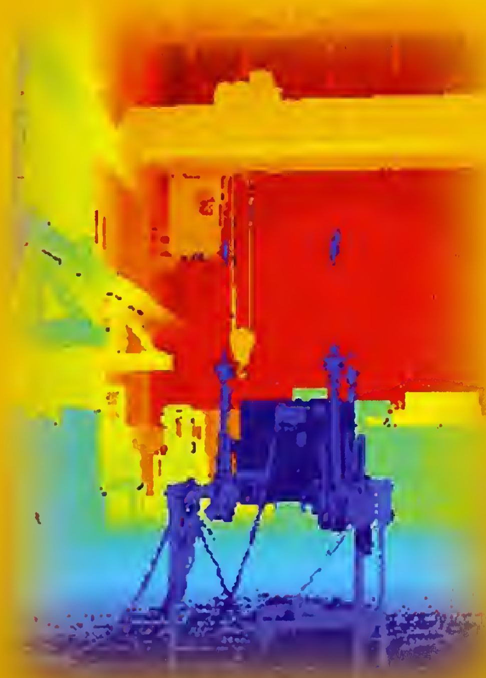
I S A R C 2 0 0 2

19th INTERNATIONAL SYMPOSIUM ON AUTOMATION AND ROBOTICS IN CONSTRUCTION



NIST
National Institute of Standards and Technology
Technology Administration
U.S. Department of Commerce

NIST SP 989

The National Institute of Standards and Technology was established in 1988 by Congress to "assist industry in the development of technology . . . needed to improve product quality, to modernize manufacturing processes, to ensure product reliability . . . and to facilitate rapid commercialization . . . of products based on new scientific discoveries."

NIST, originally founded as the National Bureau of Standards in 1901, works to strengthen U.S. industry's competitiveness; advance science and engineering; and improve public health, safety, and the environment. One of the agency's basic functions is to develop, maintain, and retain custody of the national standards of measurement, and provide the means and methods for comparing standards used in science, engineering, manufacturing, commerce, industry, and education with the standards adopted or recognized by the Federal Government.

As an agency of the U.S. Commerce Department's Technology Administration, NIST conducts basic and applied research in the physical sciences and engineering, and develops measurement techniques, test methods, standards, and related services. The Institute does generic and precompetitive work on new and advanced technologies. NIST's research facilities are located at Gaithersburg, MD 20899, and at Boulder, CO 80303. Major technical operating units and their principal activities are listed below. For more information visit the NIST Website at http://www.nist.gov, or contact the Publications and Program Inquiries Desk, 301-975-3058.

Office of the Director
- National Quality Program
- International and Academic Affairs

Technology Services
- Standards Services
- Technology Partnerships
- Measurement Services
- Information Services
- Weights and Measures

Advanced Technology Program
- Economic Assessment
- Information Technology and Applications
- Chemistry and Life Sciences
- Electronics and Photonics Technology

Manufacturing Extension Partnership Program
- Regional Programs
- National Programs
- Program Development

Electronics and Electrical Engineering Laboratory
- Microelectronics
- Law Enforcement Standards
- Electricity
- Semiconductor Electronics
- Radio-Frequency Technology[1]
- Electromagnetic Technology[1]
- Optoelectronics[1]
- Magnetic Technology[1]

Materials Science and Engineering Laboratory
- Intelligent Processing of Materials
- Ceramics
- Materials Reliability[1]
- Polymers
- Metallurgy
- NIST Center for Neutron Research

Chemical Science and Technology Laboratory
- Biotechnology
- Process Measurements
- Surface and Microanalysis Science
- Physical and Chemical Properties[2]
- Analytical Chemistry

Physics Laboratory
- Electron and Optical Physics
- Atomic Physics
- Optical Technology
- Ionizing Radiation
- Time and Frequency[1]
- Quantum Physics[1]

Manufacturing Engineering Laboratory
- Precision Engineering
- Manufacturing Metrology
- Intelligent Systems
- Fabrication Technology
- Manufacturing Systems Integration

Building and Fire Research Laboratory
- Applied Economics
- Materials and Construction Research
- Building Environment
- Fire Research

Information Technology Laboratory
- Mathematical and Computational Sciences[2]
- Advanced Network Technologies
- Computer Security
- Information Access
- Convergent Information Systems
- Information Services and Computing
- Software Diagnostics and Conformance Testing
- Statistical Engineering

---

[1] ... Boulder, CO 80303.
[2] ...ne elements at Boulder, CO.

NIST Special Publication 989

19th International Symposium on Automation and Robotics in Construction

William C. Stone, Editor

Materials and Construction Research Division,
Building and Fire Research Laboratory,
National Institute of Standards and Technology,
Gaithersburg, MD 20899-8611

September 2002

U.S. Department of Commerce
Donald L. Evans, Secretary

Technology Administration
Phillip J. Bond, Under Secretary for Technology

National Institute of Standards and Technology
Arden L. Bement, Jr., Director

National Institute of Standards and Technology Special Publication 989
Natl. Inst. Stand. Technol. Spec. Publ. 989, 557 pages (September 2002)
CODEN: NSPUE2

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 2002

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov — Phone: (202) 512-1800 — Fax: (202) 512-2250
Mail Stop SSOP, Washington, DC 20402-0001

MESSAGE FROM THE ORGANIZERS

Welcome to the 19th International Symposium on Automation and Robotics in Construction. We at the National Institute of Standards and Technology (NIST) are honored to host this prestigious conference and hope that you will find the technical sessions, demonstrations, and after-hours festivities to have made your trip to Gaithersburg, Maryland not only worthwhile, but memorable.

This year the conference has 88 papers from 20 countries. There are nine sessions including conference-traditional topics such as Design and Development of Construction Robotics; Control Systems for Construction Equipment; Construction Process Modeling and Simulation; Automated Inspection and Maintenance Management Systems; Design Practices to Facilitate Construction Automation; and Construction Information Management Systems. NIST has long held the position that to effectively bring automation to the average construction site one had to do three additional things: 1) provide affordable, ubiquitous, and powerful sensors to monitor construction status; 2) to encourage the development of standard protocols for communication of site data to the business managers, foremen, and machines that need the data; and 3) to provide compelling economic evidence to facilities owners and operators that automation is worthwhile: in short that there is a positive return on investment. We are pleased this year to present three sessions aligned with these goals: Advanced Sensing and Imaging Technologies – including LADAR, Photogrammetry, GPS, and 3D Object Recognition; Field Sensor Data and Construction Process Integration Protocols; and Economic Assessment of Automation and Robotics Technologies.

Talking about robotics, automation, autonomy, and tele-operation is one thing; seeing it and getting your hands on it in person is quite a different experience. With that in mind we have set aside a portion of the conference for a series of live demonstrations that include automated steel frame building construction; autonomous off-road vehicle navigation; tele-operative control of a remote earthmoving machine; 3D interactive visualization of 4D construction simulation; 3D laser radar scanning of construction sites; and wireless ad-hoc networks for authorization, location, and emergency notification of crafts and trades on construction sites. All of these live events will take place on the NIST campus.

Many people both within NIST and on the ISARC conference steering committee did an exceptional job of making this conference happen. We would like to especially acknowledge:

James Albus, NIST
Roger Bostelman, NIST
Robert Chapman, NIST
Geraldine Cheok, NIST
Nicholas Dagalakis, NIST
Martin Fischer, Stanford Univ.
James Garrett, Carnegie Mellon Univ
Carl Haas, Univ of Texas, Austin
Tsai Hong, NIST
Richard Jackson, FIATECH
Maris Juberts, NIST
SoonWook Kwon, Univ of Texas, Austin
H.S. Lew, NIST
Robert Lipman, NIST
Haitham Logman, Univ of Texas, Austin
Alan Lytle, NIST
Rajmohan Madhavan, NIST
Elena Messina, NIST
Richard Norcross, NIST
Mark Palmer, NIST
Frederick Proctor, NIST
Kent Reed, NIST
Kamel Saidi, NIST
Clifford Schexnayder, Arizona State
Craig Schlenoff, NIST
Michael Shneier, NIST
Iris Tommelein, UC Berkeley

for their efficient peer-reviews of the conference papers. Special thanks are due to Kathleen Kilmer, Patrice Boulanger, and Kimberly Snouffer of the NIST Conference Program Office for their dedicated and helpful assistance. The organizers also would like to extend their sincere thanks to Noel Raufaste for his extraordinarily efficient assistance with keeping all the details in line and encouraging us to meet all the deadlines.

S. Shyam Sunder
General Chair

William C. Stone
Technical Chair

Stephen Cauffman
Secretariat

ABSTRACT

This publication is the Proceedings of the 19th International Symposium on Automation and Robotics in Construction (ISARC). The symposium was held at the National Institute of Standards and Technology, Gaithersburg, Maryland during 23-25 September 2002. The Proceedings include the technical program, list of Board of Directors, Message from the Organizers, and the 88 technical papers from 20 countries authored for this international meeting.

The manuscripts were presented during nine Sessions: Construction Information Management Systems; Construction Process Modeling and Simulation; Design Practices to Facilitate Construction Automation; Design and Development of Construction Robotics; Economic Assessment of Automation and Robotics Technologies; Field Sensor Data and Construction Process Integration Protocols; Automated Inspection and Maintenance Management Systems; Control Systems for Construction Equipment; Advanced Sensing and Imaging Technologies.

KEYWORDS: Advanced sensing; automation; construction; construction equipment; construction process integration; control systems; 4D construction simulation; GPS; imaging technologies; information management; inspection; LADAR; robotics; simulation; 3D Object Recognition

CONTENTS

SESSION 8 CONTROL SYSTEMS FOR CONSTRUCTION EQUIPMENT

SESSION 9 AUTOMATED INSPECTION AND MAINTENANCE MANAGEMENT SYSTEMS

ACKNOWLEDGEMENTS

Symposium Host

The Building and Fire Research Laboratory, National Institute of Standards and Technology

Symposium Sponsor

The International Association for Automation and Robotics in Construction
Officers and Directors

Prof. Carlos Balaguer (Spain) - President
Prof. Miroslaw Skibniewski (USA) - Immediate Past-President
Prof. Thomas Bock (Germany) - President-Elect
Prof. Eugeniusz Budny (Poland)
Prof. Naruo Kano (Japan)
Prof. Ger Maas (Netherlands)
Prof. Osama Moselhi (Canada)
Prof. Derek Seward (Great Britain)
Prof. Wolfgang Poppy (Germany)
Dr. M.T. Wang (Taiwan)
Dr. Ronie Navon (Israel)
Dr. Carl Haas (USA)
Dr. N. Pai (Taiwan)
Mr. Tetsuji Yoshida (Japan)
Mr. Par Ahman (Sweden)
Mr. Francois Peyret (France)
Mr. Hannu Koski (Finland)
Mr. Johnatan O'Brien (Australia)
Mr. Fiorenzo Malaguti (Italy)

Symposium Co-Sponsors

Automation in Construction, An International Journal, Elsevier
Center for Integrated Facilities Engineering, Stanford University
Construction Engineering and Management Division, Purdue University
Construction Industry Institute
Construction Institute, American Society of Civil Engineers
Continental Automated Building Association
Del E. Webb School of Construction, Arizona State University
FIATECH Consortium
Institute for Research in Construction, National Research Council, Canada
International Alliance for Interoperability, North America
Lean Construction Institute

19th ISARC Steering Committee

Dr. Sherif Barakat, Institute for Research in Construction, National Research Council of Canada
Dr. Leonhard E. Bernold, North Carolina State University, Center for Construction Technology and Integration
Mr. Stephen A. Cauffman, Vice Chair (Secretariat), National Institute of Standards and Technology
Lt. Gen. Kenneth E. Eickmann (USAF, Retd.), Construction Industry Institute
Dr. John M. Evans, National Institute of Standards and Technology
Dr. Martin Fisher, Stanford University, Center for Integrated Facilities Engineering
Ms. Karen M. Furlani, National Institute of Standards and Technology
Dr. James H Garrett, Jr., Carnegie Mellon University
Dr. Francois Grobler, U.S. Army Engineer Research and Development Center, International Alliance for Interoperability - North America
Dr. Carl T. Haas, University of Texas at Austin, Construction Automation Laboratory
Dr. Miriam Heller, National Science Foundation
Dr. Richard H.F. Jackson, FIATECH Consortium
Dr. Ronie Navon, Technion - Israel Institute of Technology, Construction Automation Laboratory
Dr. Francois Peyret, Laboratoire Central des Ponts et Chaussees, France
Dr. Jeffrey S. Russell, University of Wisconsin-Madison, Construction Institute/American Society of Civil Engineers
Dr. Clifford J. Schexnayder, Arizona State University, Del E. Webb School of Construction
Dr. Miroslaw Skibniewski, Purdue University, International Association for Automation and Robotics in Construction
Dr. William C. Stone, Vice Chair (Proceedings Editor), National Institute of Standards and Technology
Dr. S. Shyam Sunder, Chair, National Institute of Standards and Technology
Dr. Iris D. Tommelein, University of California at Berkeley, Lean Construction Institute
Dr. Richard L. Tucker, University of Texas at Austin, Center for Construction Industry Studies
Dr. John G. Voeller, Black & Veatch Holding Company

SESSION 1

CONSTRUCTION INFORMATION MANAGEMENT SYSTEMS

Collaborative Fire Protection Engineering Based on Software Agents

by

Uwe Rueppel[1]

ABSTRACT: This paper presents a new approach with software agents to support the task of preventive fire protection engineering. Agents are stationary or mobile software components to support the user in his planning tasks according to their specific design. Agents can act autonomously within their environment, which means that they can solve problems without the need of control through a human being or another system. In order to react to modifications of the environment agents observe their environment continuously and react adequately to modifications. To reach a result it is necessary for autonomous agents to react particularly in a goal-oriented way. The purpose of an agent system in the area of fire protection engineering is to organize a distributed problem solution by many agents representing adaptive experts. Using available object-oriented, net-enabled models of standard applications and based on the technology of the software agents, a platform for collaborative fire protection engineering is designed. Local service-agents provide necessary information in a collaboration platform. By communication between these local service-agents and mobile verifying-agents the appropriate information can be processed to verify the fire protection model of a building.

KEYWORDS: Collaboration platform; fire protection engineering; information technology; software agents;

1. INTRODUCTION

Fire protection engineering is a part of the building design process in civil engineering and consists of the cooperation of experts in multiple disciplines. Consistent cooperation of engineers in different fields is the requirement for high product quality, for optimal and short development with a minimum of investment costs. The development process in building engineering is strongly assigned by creating unique buildings. Software methods and tools in building engineering do not sufficiently support the communication, collaboration and control processes, especially in the area of fire protection engineering. Most software solutions are based on proprietary solutions with their specific technical model. Collaboration is often implemented as a centralized client-server architecture. Model information in these client-server architectures are static with the consequence that every completion of a new model requires the definition of new interfaces and implementations in every module [1].

[1] Professor for Computer Science in Civil Engineering
Darmstadt University of Technology
Petersenstrasse 13, 64287 Darmstadt, Germany
Tel.: +49 (0) 6151 16 34 44
Fax: +49 (0) 6151 16 55 52
rueppel@iib.tu-darmstadt.de

2. Fire Protection Engineering

In addition to the described problems in collaboration, a great number of codes and regulations for fire protection models of buildings have to be evaluated. In Germany every state has different fire protection codes and a great number of special building regulations (for multi-storey buildings, hospitals, etc.) (see Figure 1). To consider these complex standards completely and without errors during the design process is a challenge in collaborative fire protection engineering. Fire protection models therefore are often inconsistent.

German codes define the fire protection needs overall. There are different parts of fire protection engineering:

- Organizational;
- Operational;
- Defensive; and
- preventive.

The focus of the presented approach is the preventive fire protection engineering. This part has to be recognized during the whole planning and construction phase and is the basis for the defensive fire protection. The goals in preventive fire protection engineering are:

- prevention of fire;
- containing the fire using fire compartments;
- developing an infrastructure for extinguishing the fire (fire extinguisher, sprinkler etc.)
- and developing an infrastructure for rescue.

As it is shown in Figure 2, these preventive fire protection elements are represented by or have strong relationships to building elements.

Normally the fire protection model is developed by architects as a part of their building design. Based on this draft, other planners have to complete the detailed planning and advance the fire protection model continuously (see Figure 3). That means, that for example a structural engineer finds restrictions due to fire resistance demands in the choice and design of the structural elements. The technical services for facilities have to be designed according to the fire protection model, so that planned rescue routes are free of smoke from flammable components [2]. After some great fire disasters in Europe, the compliance and inspection of fire protection rules are taken very seriously. With regard to the devastating consequences to people living and working in buildings, it is necessary to include the fire protection concepts very early in the building design process being conducted by the all specific planners involved.

In order to support the disciplinary processes of the fire protection model evolution in a distributed environment, the integration of all participants in the entire planning process is necessary. Therefore data, methods and knowledge of the partial models are to be coupled for collaboration, while taking into consideration that the planners are spatially distributed. The concept of collaborative work of distributed planners must support sequential, parallel and iterative planning processes within an integrated information platform.

3. Software Agents

An agent is one that is authorized to act for another. Agents possess the characteristics of delegacy, competency and proactive behaviour. Delegacy means the discretionary authority to autonomously act on behalf of the client. Actions include making decisions, committing resources, and performing tasks. Competency is the capability to effectively manipulate the problem domain environment to accomplish the prerequisite tasks. Competency includes specialized communication proficiency. Proactive behaviour is the ability to adapt behaviour to optimize performance in an often non-stationary environment in responsive pursuit of the goals of the client.

A software agent is an artificial agent which operates in a software environment. Software environments include operating systems, computer applications, databases, networks and virtual domains. Delegacy for software agents means that they stay resident or persistent as background processes after being launched. By making decisions and acting on their environment independently, software agents reduce human workload by generally only interacting with their end-clients when it is time to deliver results. Competency within a software environment requires knowledge of the specific communication protocols of the domain. Protocols such as SQL for databases, HTTP for the WWW and API calls for operating systems must be implemented into the software agents. Proactive behaviour for software agents is generally limited to providing control options and the generation of status reports that require human review.

Mobile Agents, also known as travelling agents, will shuttle their code and state among resources. This often improves performance by moving the agents to where the data reside instead of moving the data to where the agents reside. The alternative typical operation involves a client-server model. In this case, the agent, in the role of the client, requests that the server transmit volumes of data back to the agent to be analyzed. Oftentimes the data must be returned by the agent to the server in a processed form. With mobile agents load-balancing can be achieved by distributing agents over a finite number of computational resources. Some mobile agents are self-distributing, seeking and moving to agent platforms that can offer the higher computational resources at lower costs. Further a complex problem solution can be achieved by dividing the problems into less complex sub-problems addressed by collaborating agents.

Collaborative agents interact with each other to share information or barter for specialized services to effect a synergism. While each agent may uniquely speak the protocol of a particular operating environment, they generally share a common interface language which enables them to request specialized services from other agents as required. This means that agents have to communicate [3]. The communication and cooperation with other agents, machines or people is realized by specific languages as KQML [4] [5], ACL [6] or Tcl [7].

4. Fire Protection Engineering with Software Agents

These qualities of agent systems are utilized for the evolution of an integrated fire protection model. The information of the fire protection model is determined from two different sources. On the one hand, fire protection relevant rules must be accessible to the system. On the other hand, the information of building product models must be available. The fire protection model must be adapted to the different views of the planners within the collaborative building design and relevant information must be evaluated within the agent-based platform [8]. The system architecture of the collaboration platform for preventive fire protection engineering is shown in Figure 4.

The determination and access of the fire protection information are designed as a mobile verifying-agent. The architect designs a fire protection model on the basis of his building design in his working environment. He defines the elements, relevant. for preventive fire protection. A fire protection application for AutoCAD ADT was developed for this purpose. On completion of a design, the architect publishes the fire protection model on his agent-platform. For this purpose a fire protection- XML-interface for AutoCAD ADT was developed. The verifying-agents are then available on the server with the fire protection model in order to support the collaboration of all other planners. On the one hand, every planner can retrieve the fire protection model from the server in the form of mobile agents and check his own models for consistency with the fire protection model. The planner is supported by the agent in the form of a check list. A fire protection inspection can occur at the time the agent gets all information needed. On the other hand, the architect can verify the released technical information of the planners for consistency with the fire protection model by mobile verifying-agents. For this purpose he defines the run places for the verifying-agents representing the cooperation partners. The mobile verifying-agent determines the necessary information from every cooperation partner. This can come about automatically, when on the one hand a link is already defined between the corresponding technical model (e.g. building services model or structural model [8]) and the fire protection model. On the other hand, a communication of the responsible planner can be initiated to link the information.

Each planner makes his model information available to other participants in the design process through a local service-agent. This service-agent has control over the specific model information and can supply the verification. The protection of the planner specific information is guaranteed by encapsulating it within the interface of the service-agent. Other users cannot manipulate the information since the encapsulation is controlled completely by the local service-agent.

5. CONCLUSIONS

This paper described how the technology of software agents enables the integration of different technical models to support collaborative fire protection engineering. A new methodology was developed to verify fire protection models in a distributed environment by means of software agents. This approach integrates the information of building design, structure and technical services with knowledge and rules of German fire protection codes (see Figure 4). The communication between the integrated fire protection model and local technical models of the planners is integrated into the system. The concept allows the integration of fire protection regulations into the planning processes during early design stages. The new software approach enables collaborative engineering with fire protection models and leads to an increase of quality and consistency in the resulting preventive fire protection design system [8] [9].

6. REFERENCES

[1] Rueppel, U.; Meissner, U.: Cooperative Structural Engineering in Distributed Systems. In: Proceedings of the 8th International Conference on Computing in Civil & Building Engineering, ISBN 0-7844-0513-1, P. 504-509, Stanford University, USA 2000.

[2] Loebbert, A., Pohl, K.D. and Thomas, K.-W.: "Brandschutzplanung fuer Architekten und Ingenieure (Fire Protection Engineering for Architects and Engineers)", Verlag Rudolf Mueller, Koeln, Germany, 2000.

[3] Brenner, W., Zarnekow, R. and Wittig, H.: "Intelligente Softwareagenten (Intelligent Software Agents)", Springer Verlag, Berlin, Germany, 1998.

[4] Finin, T., Fritzson, R., McKay, D. and McEntire, R.: "KQML as an Agent Communication Language", In: Proceedings of the third International Conference on Information and Knowledge Management (CIKM'94), ACM Press, USA 1994.

[5] Labrou, T. and Finin, T.: "A Proposal for a new KQML Specification", In: Technical Report TR CS-97-03, University of Maryland, USA, 1997.

[6] Fipa: "ACL - Agent Communication Language", Foundation for Intelligent Physical Agents (FIPA) specifications homepage. Available online at http://www.fipa.org, 06-2002.

[7] Oustershout, J.K.: "Tcl and the Tk Toolkit", Addison Wesley, New York, USA, 1994.

[8] Meissner, U.; Rueppel, U.; Petersen, M.: Cooperative Management of Product Models in the Planning Process of Structural Engineering. In: Proceedings of the International Conference for Computational Engineering Science (ICES2K), ISBN 0-9657001-3-5, P. 240 - 245, Los Angeles, USA 2000.

[9] Rueppel, U.; Theiss, M.: Cooperative Developing of Building Projects in Computer Networks. In: Proceedings of the International Conference for Computational Engineering Science (ICES2K), ISBN 0-9657001-3-5, P. 1182- 1187, Los Angeles, USA 2000.

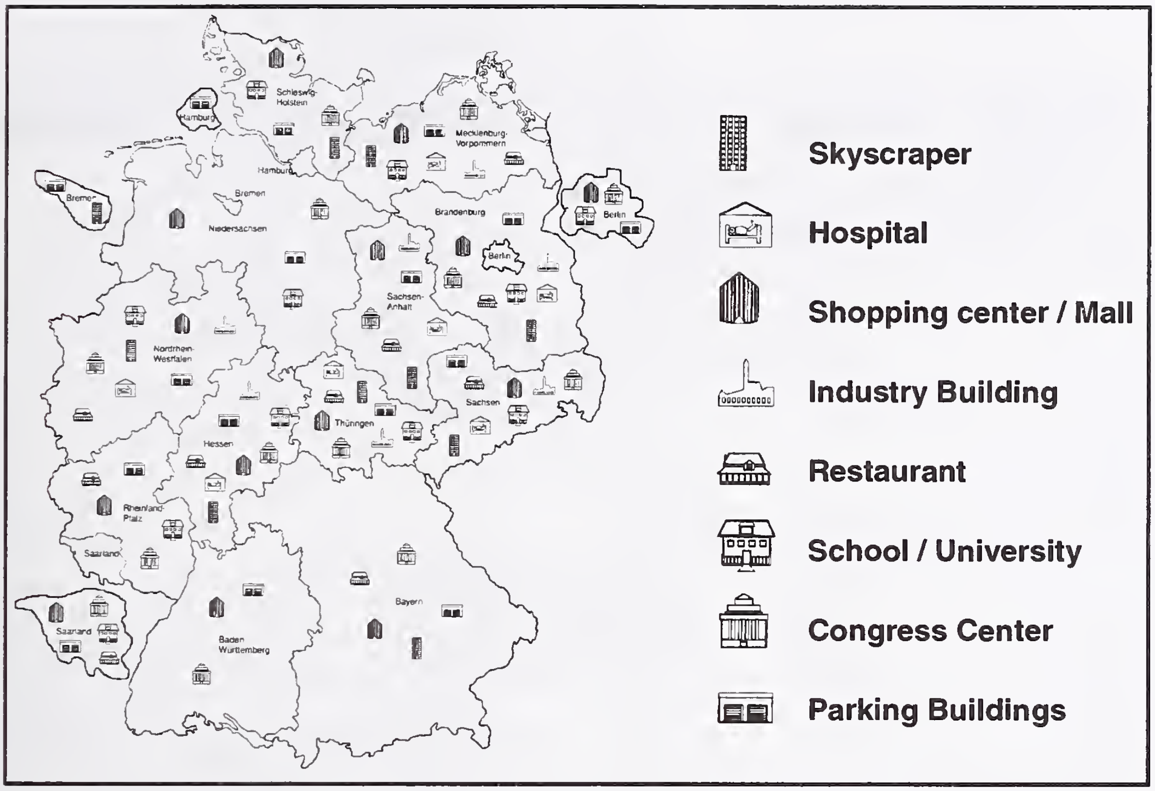


Figure 1: Fire Protection Regulations in Germany

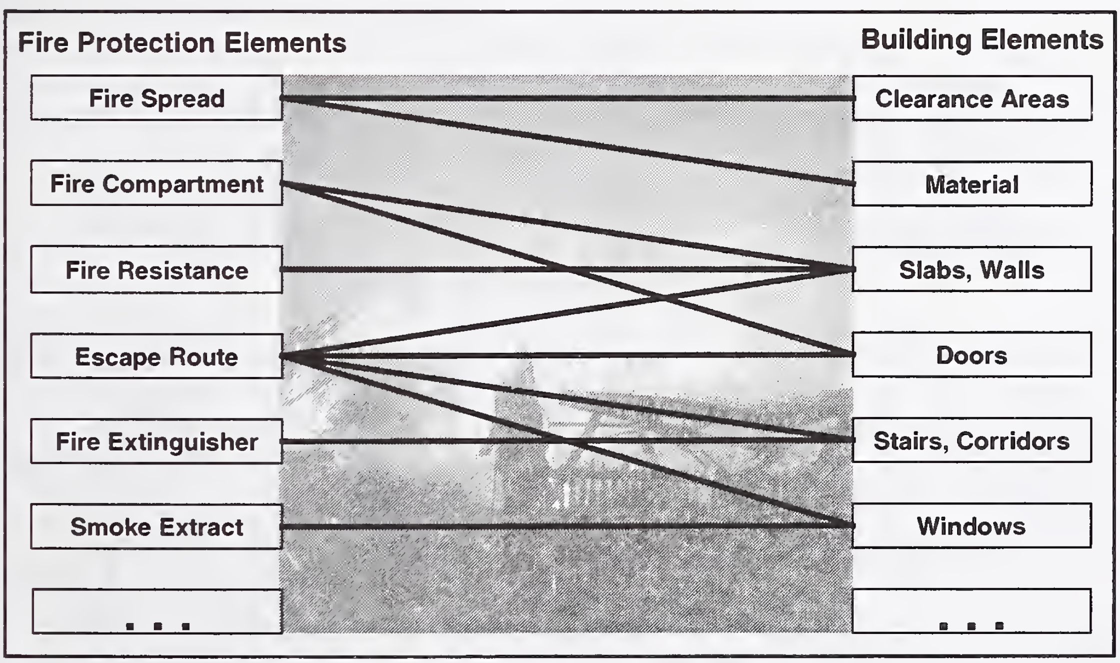


Figure 2: Fire Protection and Building Engineering

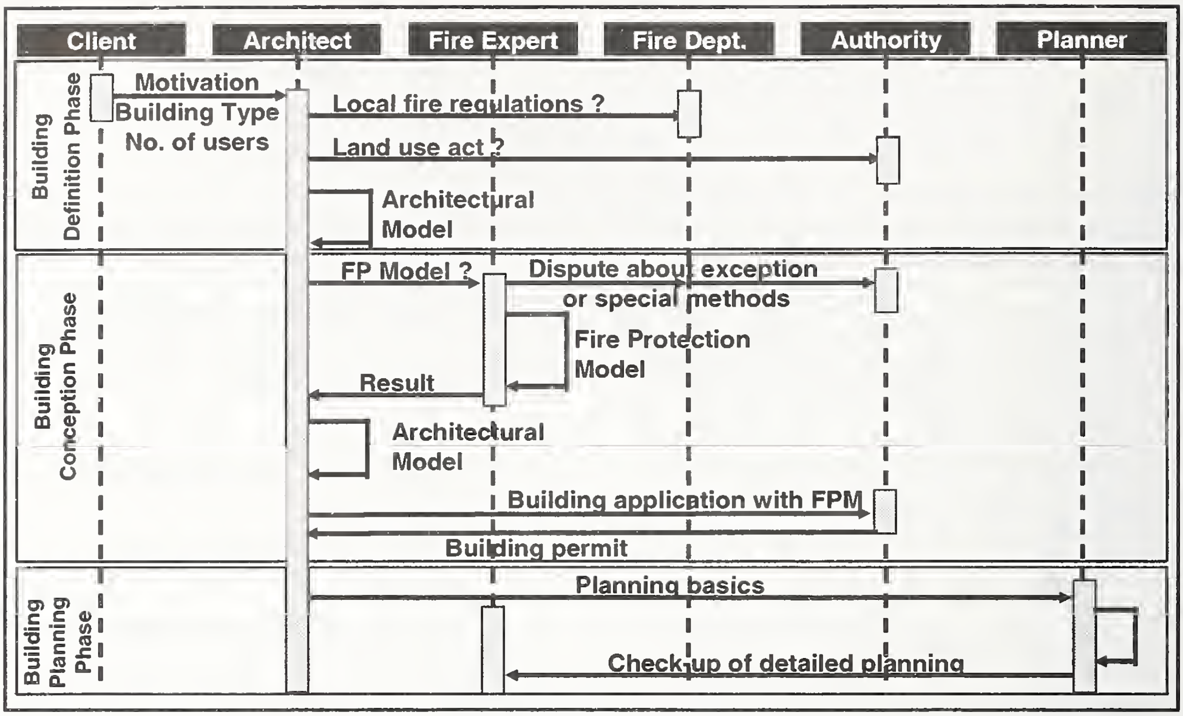


Figure 3: UML Diagram for Fire Protection Planning

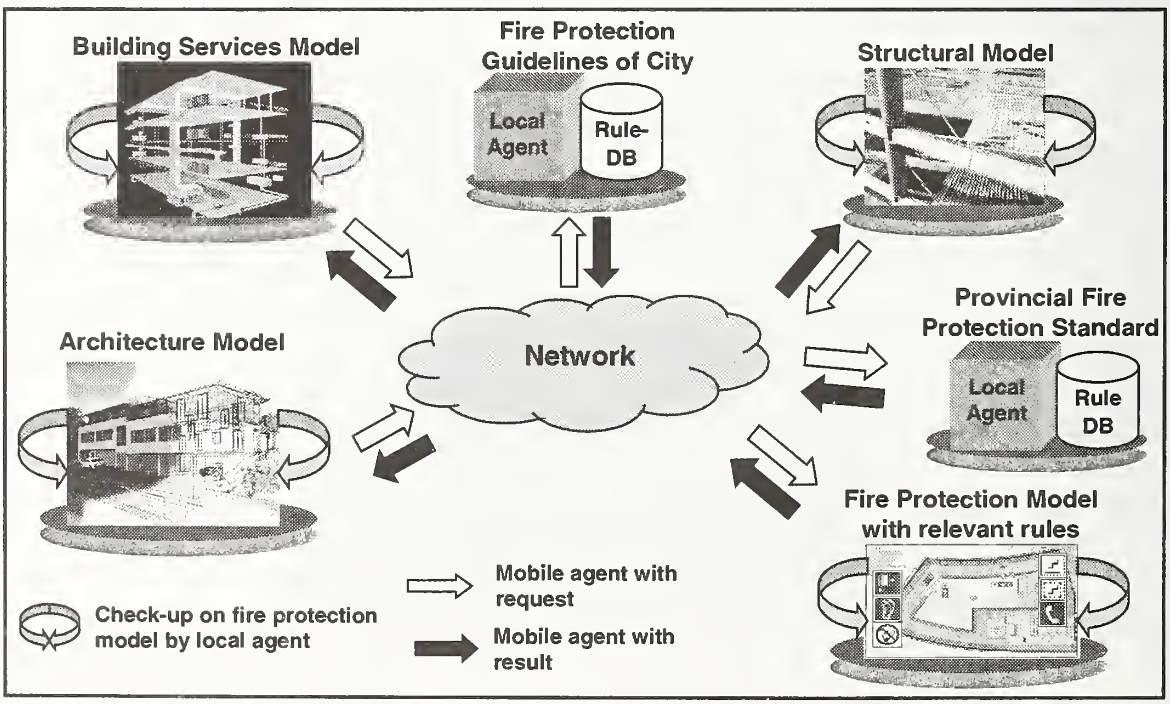


Figure 4: Platform Architecture for Collaborative Fire Protection Engineering

Strategies for Realizing the Benefits of 3D Integrated Modeling of Buildings for the AEC Industry

by

Chuck Eastman*, Rafael Sacks**, Ghang Lee***

ABSTRACT: We explore the reasons why advanced information technology applications have not been adopted by the AEC industry. Within the industry, however, some sectors have made significant moves toward adoption of advanced IT. By examining these areas of success, we propose a framework for the incremental conversion of the AEC industry to fully incorporate advanced IT.

KEYWORDS: knowledge-based applications, building models, information technology innovation, construction automation.

1. INTRODUCTION

The business models in each industrial domain provide different contexts for information technology innovation. Here, we focus on the building design and construction part of the Architecture, Engineering and Construction (AEC) industry. The building industry custom designs, engineers and constructs projects on varied construction sites. The components are either procured off-the-shelf, made-to-order or fabricated on-site. A variety of design and engineering specialists collaborate on a project, each bringing specific expertise and employing specialized software. In the US, the value of all construction work is 845 billion dollars (Economic Census, 2000), making it one of the largest industries in the nation.

Digital knowledge-based models of buildings have been advocated as the base representation for work in the construction industry for over twenty-five years (Eastman, 1999, Chapter 2). Here, we use "knowledge-based building models" as a shorthand reference for the integrated use of several technologies: (1) multiple heterogeneous computer applications allowing three dimensional representation, design, analysis and management of the systems and components that make up the proposed building, all incorporating significant domain knowledge, (2) the backend integration of workflow using one or more building product models; (3) associated Internet links to material supplies, delivery planning, operational controls and other such services needed for the building's procurement, fabrication and operation, and (4) expanded automation of all aspects of work, including design automation, automation of built-to-order components, and automation of on-site fabrication. We consider these technologies together to define current best practice use of advanced Information Technology (IT).

Although serious building model efforts have been carried on for ten years (e.g., CIMsteel – CIS/2 2001), national and international efforts in moving the AEC industries toward the use of advanced information technologies has shown little progress. Virtually all buildings designed today still rely on two-dimensional (2D) drawings as their principal representation. This is not to say that there has been no movement; rather, the particular conditions in the construction industry have not allowed it to proceed as fast as other product-oriented domains, such as manufacturing, electronics, aerospace and other industrial areas. It has also not even progressed as fast as art or entertainment industries, such as television, motion pictures, or music production.

This paper reviews why there has been so little movement and outlines a strategy for the incremental conversion of the AEC industry to the use of advanced IT applications. We build upon two examples, involving the structural

* Professor and Director of the College of Architecture Ph.D. Program, and Professor of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0155; chuck.eastman@arch.gatech.edu

** Research Faculty, College of Architecture Ph.D. Program, Georgia Institute of Technology, Atlanta, GA 30332-0155; rafael.sacks@arch.gatech.edu

*** Ph.D. research Assistant, College of Architecture Ph.D. Program, Georgia Institute of Technology, Atlanta, GA 30332-0155; ghang.lee@arch.gatech.edu

steel and precast concrete sectors of the building industry. The issues and strategy may be useful for other industries, beyond AEC.

2. WHY THE CONSTRUCTION INDUSTRY HAS LAGGED

The commonly held assumption why building design, engineering and construction have not utilized knowledge-based building models is due to *fragmentation*. The construction industry has traditionally been craftsman-oriented, with many small participants. There are 656,000 construction-related firms in the United States alone. Of these, over 62 percent have fewer than 4 employees, while less than 1 percent have more than 100 (ICAF, 2000). Each firm carries out only a few steps of an increasingly complex process. No one organization has as much as five percent of construction dollars. These many organizations come together to execute construction projects in a very competitive setting, with each project typically involving a change of participants.

Within this fragmented setting, 2D architectural and engineering drawings have evolved over hundreds of years as the basic representation and physical artifact used by all construction industry participants. Business practices of all construction-related firms, financial and insurance institutions, standards and codes and reference materials are organized around 2D drawings and associated formatting conventions. However, it is generally recognized that it is not possible to verify the consistency of 2D drawings or to derive all the needed geometry for design and production automation. These are only possible with 3D modeling.

The very broad embedding of 2D drawings, combined with the fragmentation of the participants, is assumed to result in the situation where no single participant and no single project have enough economic impact to justify the investment – in technology purchases, in learning and in changing business practices – to convert to knowledge-based building models. No organization has the power to force conversion of the others (seemingly, not even the US government).

The potential economic benefits of knowledge-based building models are varied in nature: some occur in isolated construction processes (e.g. CNC production of made-to-order components), while others depend on improving the process as a whole (e.g. reduction of re-work resulting from errors). The motivations cited in most research have emphasized long-term system benefits without identifying measurable short-term payoffs (Moreau and Back 2000, Griffis and Sturts 2000). Because no one party to the building process seems to benefit, but the process as a whole improves, it is argued that the owners and clients of buildings are the beneficiaries and should financially support this IT transition. However, no effective business model has been implemented based on this premise.

We accept these conditions, in general, and conclude that *IT technology innovation must evolve based on local benefits*, not industry-wide ones.

3. SECTOR-LEVEL ADOPTION OF ADVANCED IT

Given the complex conditions cited above, it is apparent why current strategies have not moved the construction industry toward knowledge-based building models; there is no one business entity or financial incentive large enough to move the industry as a whole. (Griffis et al. 1995) and production (FIATECH 2001) indicates that the potential exists for significant process improvements in construction itself, and that these process improvements require IT support.

Despite this state of affairs, certain sectors of the construction industry have made strong steps to implement knowledge-based building models within their sector. We review two: structural steel and precast concrete.

Structural Steel

Structural steel fabrication and erection is a significant sector within the construction industry, with 8.5 billion dollars of production (Economic Census, 2001) and half a million workers (US Dept. of Labor 1999). Steel is used in a wide range of structures, including buildings, industrial facilities and process plants. Engineering companies, steel fabricators and structural engineers, as well as architects, are involved in aspects of the building structural steel sector. The processes in this sector include designing, detail engineering, prefabricating of pieces, shipping and erection of the pieces on the construction site.

Steel design and fabrication involves a significant amount of engineering that can benefit from knowledge-based design automation. Industry experts suggest that steel fabrication, which is dependent on engineering design, increases the value of raw steel sections by a factor of 2.5. The sector is supported by more than 30 commercial computer applications, providing computational services for design, analysis, detailing, bills of material and material tracking, scheduling and fabrication management (ERP) programs. A significant number of these applications are based on 3D modeling of the structure and pieces. The sector is also supported by a variety of production automation equipment -- for cutting, drilling, welding and other operations. Steel fabricators take different roles within the construction process. While the most common is as a subcontractor, they sometimes serve as general contractor; in some cases, steel fabricators have embraced design-build and offer package services.

In the mid-1990s, the American Institute of Steel Construction (AISC), the main industry sector association, adopted the Structural Design Neutral File (SDNF) as a way to exchange data between the various applications. When limitations in SDNF were found that restricted its use, the AISC initiated an Electronic Data Interface (EDI) initiative to select a more robust and complete data representation for integration and exchange of structural steel data. In late 1999, they adopted CIMsteel Integration Standard, Version Two (CIS/2), developed at Leeds University, as part of a European Union automation effort. Major efforts have been directed toward its deployment, and at this time, ten applications have developed working CIS/2 interfaces. Several building projects have now been built using CIS/2 and it is coming into wide use within this industry sector (CIS/2 2001).

Precast Concrete

The precast concrete sector -- prefabricating concrete products off-site, often prestressed, then shipping and erecting them – is a fairly young sector of the construction industry. Compared to structural steel, the precast concrete sector in the US is smaller, with 2.6 billion dollars of annual work (Economic Census, 2001- non-building products omitted), about 200,000 employees. Industry experts suggest that the fabricated value of precast products is on the order of 6.0 times the raw material costs, again, with a large engineering component. In Europe, this industry sector is much larger; its share is above 15% while in the US, it deals with 1.2% of construction (Sacks et al. 2002). Like structural steel, careful design and coordination with the rest of the building design is required for the prefabricated pieces to fit correctly when erected. While the level of automation in the US is generally low, it is more advanced in Europe, providing both mature technology to draw upon and examples of its use. Precast concrete fabricators assume varying roles within the construction process: while the most common is as a sub-contractor, a significant number of precast fabricators provide design-build package services for specific structures, such as parking garages.

In contrast to steel, the precast concrete industry sector has been supported by only a small number of commercial software applications, with small sales volumes. Many of these have been custom-programmed and paid for by individual precast producers. Some software has been funded by industry associations, primarily the Precast Concrete Institute. A significant percentage of precast fabricators have implemented Enterprise Resource Planning (ERP) systems.

In 2000, the North American precast industry initiated an information technology initiative (see http:// usa.arch.gatech.edu/pci/), using a limited liability corporation called the Precast Concrete Software Consortium (PCSC). The first step of the PCSC was to undertake careful process modeling by the member companies, to gain understanding of their current workflow and identify opportunities for IT (Eastman et al. 2002). They then developed a plan that included specification for design and engineering software of precast concrete building assemblies and also individual pieces (Eastman et al. 2001) and for the development of a precast concrete building model. The software specification defines a knowledge-based CAD application for the precast concrete industry. The specification was completed in

December 2001 and distributed to twenty-six software organizations; twelve proposals were submitted. Today, they are negotiating with two candidate companies to develop and implement software. Definition of the precast concrete building model was begun in the Spring of 2002.

It is clear that the internal conditions within these two sectors of the building industry were appropriate to make significant investments to develop and implement knowledge-based building models, as we have defined that term. The benefits were not industry-based, but rather for their own sector and companies. The authors propose that given these initiatives in IT and their different situations, it may be possible to generalize and posit a framework for IT development that lays out the necessary conditions for IT innovation. If such a framework was developed, it would allow us to identify necessary conditions for particular types of IT innovation to be made, and to recognize whether an industry sector was "ready" (or not) for various innovations leading to knowledge-rich building models.

Below, we propose such a framework of technology innovation within a complexly structured industry, focused on IT innovation in construction.

4. A FRAMEWORK FOR IT INNOVATION IN THE CONSTRUCTION INDUSTRY

The adoption of knowledge-based building models by companies in the construction industry is obviously dependent on many factors. We have identified seven conditions, organized hierarchically:

<u>Pre-conditions;</u> only if these conditions exist, or are perceived to exist by industry leaders, will there be motivation to take risks and expend intellectual development effort to undertake industry sector change:

(1) an economic situation with a significant value-added component to the sector's activities, allowing capital and knowledge-based investments;

(2) perceived benefits from technology innovation, in possible gains in market share, in capturing profits of an outside but closely related sector, or other business incentives;

<u>Leveraging conditions;</u> these conditions provide the use benefits of knowledge-based building model technology:

(3) the availability of automation technologies that require rich digital project data; the automation may be at the production level, supporting computer controlled machining, welding, concrete mixing, conveyors, finishing, etc.; it may be at the erection planning or erection level; it may be at the analysis and simulation level;

(4) computer integration in the internal business environment, which would benefit from integrating project level automation with enterprise level data management, for scheduling, procurement, manpower planning, etc. Many of the capabilities are provided by ERP systems;

(5) computer integration of aspects of the external business environment; this includes building code checking, web-based bidding for services, web-based procurement, web-based project management or other business activities with outside organizations;

<u>Information generation conditions</u>; knowledge-based data throughout the industry sector must be generated by specific applications:

(6) the availability of knowledge-based 3D modeling CAD software, capable of providing the rich digital data needed to support design or production automation within that industry sector; development of such software requires embedding much of the codified engineering knowledge in that sector into the CAD software, automating conventional design practices; data-rich CAD also provides the detailed engineering data used for production, purchasing, scheduling and other activities;

<u>Information integration and exchange;</u> the final step, in this view, allows the knowledge-based data to be fully utilized throughout the industry sector, for automation, enterprise integration and web-based e-commerce:

(7) development of a product data model for the industry sector, that supports the integration of the above aspects of the business with enterprise and web-based applications.

We consider these conditions to define a cascade or waterfall set of conditions, where previous conditions must be met before others below can be successfully implemented. These conditions are being met in the structural steel industry, and are rapidly developing in the precast concrete industry.

5. CURRENT EFFORTS TOWARD IT INNOVATION IN THE AEC INDUSTRY

Most of the conditions outlined above have not been met in the architectural or general contracting sectors. Yet, the majority of research in the building construction product-modeling arena has focused on integration at the front-end architectural design part of the process, because these activities are the primary information generators. The current major effort in developing a knowledge-based building model is centered around the International Alliance for Interoperability (IAI) and its Industry Foundation Class (IFC) building model (IFC 2001). In the IAI, a wide range of construction industry organizations have come together to fund the development of the IFCs.

However, architects and engineers have been expected to make equally significant investments in adapting their business practices to new software tools required of the IFC integration technology. Effective implementation of large IT systems requires not only vision and investment, but also significant process changes, which imply organizational restructuring (Egan 1998). While design activities have a large value-added component, architects and engineers receive a small and relatively fixed percentage of the construction dollar. No effective business plan has been implemented to pay for the additional work that populating a building model would require. Also, personnel issues and training often stand at the forefront of such restructuring.

Currently, only a few architectural or building design software applications exist that claim to be able to fully populate a building model. No set of applications, as of 2001, was able to support all the design detailing typically involved in architectural practice. Thus the software needed to take advantage of a building model has been lacking.

On the software side, the large software companies whose products are widely used in the AEC industry stand to reap relatively small and uncertain benefits from developing new generations of software technology. Such changes may even result in their losing market share when new selection criteria and lack of track records in the new functionality open the market to new competitors. On the other side, small software companies that are naturally interested in displacing the established ones, do not have the financial backing needed to educate designers to new technologies and new ways of doing design.

Because there have been few serious efforts at deployment, many technical problems have not been resolved (Amor and Faraj 2001, Eastman and Augenbroe 1998). Among the issues needing further work are: integrating applications and data exchange with productive workflows, system architectures to support integrated data models, and data management and integrity issues where many organizations and individuals contribute to a building project, with various levels of involvement and over differing segments of a project life-span. Hurdles also exist in software able to deal with the datasets associated with large, complex production-oriented 3D models. Even medium sized projects are likely to require models with on the order of a half a million objects. These issues are among those that have stifled taking advantage of recent developments regarding web-based project coordination.

While there are significant analysis and simulation applications that could be integrated, their benefits have not been viewed as significant by architectural practitioners (as evidenced by current investment). Benefits from integration with enterprise-level computer systems in architectural firms appears not to be significant. One under-developed potential benefit of knowledge-based building modeling for architects is for design automation support of design development and contract drawing preparation.

Similarly, contractor services have not been computerized beyond project estimating and billing, often because of the claim that none of their subcontractors are ready. Again, strong knowledge-based CAD systems have not been available for either industry sector.

However, the framework above suggests that different sectors of the downstream building production process, certainly including precast concrete, structural steel, but also possibly curtainwall and window wall fabricators, elevator and vertical circulation systems, HVAC systems, and others, may have better business reasons to implement advanced IT. If this happens, then the context for general contractors will change, leading to opportunities for them also to adopt advanced IT.

6. SUMMARY

We have developed a simple framework based on the above seven issues. The framework suggests a sequence of undertakings within each AEC business sector, based on the current situation of that sector. It can be used to identify, for any AEC industry segment, the economically beneficial IT development steps and the potential benefits of those changes. It avoids predicating computer integration on any single design, construction or IT technology, in favor of a systems approach, using contextually based implementation criteria.

We are using this model to evaluate AEC industry business sectors, with two goals: measuring their potential to derive benefit from knowledge-based building modeling; and proposing incremental strategies for achieving the benefits of advanced IT, with each step providing distinct payoffs. Assuming that additional industry sectors will follow the lead of the steel and precast concrete sectors in incorporating 3D modeling in their business practice, multiple critical masses may be reached, leading larger segments of the construction industry to move to integrated IT. These may percolate through the whole AEC industry, which will eventually make the transition to integrated 3D modeling.

7. REFERENCES

Amor, R., and Faraj, I., (2001). *Misconceptions about Integrated Project Databases*, Information Technology in Construction Journal (ITCON), http://www.itcon.org/2001/5/, ISSN 1400-6529, Vol. 6, pp.57-68.

CIS/2 (2001), CIMsteel2 http://dcom.arch.gatech.edu/aisc.

Eastman, C.M., Lee, G., and Sacks, R., (2002). "A New Formal and Analytical Approach to Modeling Engineering Project Information Processes", *CIB W78 Conference,* Aarhus, Denmark, June 2002.

Eastman, C.M., Sacks, R., and Lee G. (2001). "Software Specification for a Precast Concrete Design and Engineering Software Platform", *PCSC Research Report*, Georgia Institute of Technology, Atlanta GA USA.

Eastman, C.M., (1999). *Building Product Models*, CRC Press, Boca Raton, FL, USA.

Economic Census (2001). *Manufacturing Subject Series for 1997*, Doc. # EC97M31S-PS, U.S. Dept. of Commerce, Economics and Statistics Division, October, 2001.

Egan J. (1998). 'Rethinking Construction', UK Govt. Department of Trade and Industry, http://www.dti.gov.uk/construction/rethink

Economic Census, (2000) Industry Summary, 1997 Economic Census, Document EC97C23-IS, U.S. Dept. of Commerce, Economics and Statistics Division, January, 2000

FIATECH (2001). "Economic Impact of Imperfect Interoperability" FIATECH Projects http://www.fiatech.org/projects/eiii.htm

Griffis, F.H., Hogan, D. and Li, W. (1995). "An Analysis of the Impacts of Using Three-Dimensional Computer Models in the Management of Construction" *Construction Industry Institute Report RR106-11*

Griffis, F.H. and Sturts, C. (2000). "Three-Dimensional Computer Models and the Fully Integrated and Automated Project Process for the Management of Construction" *Construction Industry Institute Report RR152-11*

ICAF, Industrial College of the Armed Forces, (2000). *Construction*, available at: http://www.ndu.edu/ndu/icaf/industry/construction/construction.htm

IFC (2001) "International Alliance for Interoperability", http://www.iai.org Industry Foundation Classes IFC Release 2.X.

Moreau, K.A. and Back, W.E. (2000),"Improving the design process with information management" *Automation in Construction,* 10, pp. 127–140

Sacks, R., Eastman, C.M., and Lee, G., (2002). "Process Model Perspectives On Management and Engineering Procedures in the Precast/Prestressed Concrete Industry", submitted to the *ASCE Journal of Construction Engineering and Management*, May.

US Dept. of Labor (1999). National Industry-Specific Occupational Employment and Wage Estimates (1999). SIC 327 - Concrete, Gypsum, and Plaster Products, 1999, Bureau of Labor Statistics, US Dept. of Labor.

THE ROLE OF THE CIMSTEEL INTEGRATION STANDARDS IN AUTOMATING THE ERECTION AND SURVEYING OF STRUCTURAL STEELWORK

Kent A. Reed
Leader, Computer Integrated Construction Group
Building Environment Division
Building and Fire Research Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8630
http://cic.nist.gov
kent.reed@nist.gov

ABSTRACT: The second release of the CIMsteel Integration Standards (CIS/2) has been endorsed by the American Institute of Steel Construction as the standard for the electronic exchange of structural steel project information for the North American steel design and construction industry. Derived from the deliverables and experiences of the Pan-European Eureka CIMsteel Project and published by the Steel Construction Institute in the UK, the CIS/2 was developed with a life-cycle view of structural steelwork information in mind. However, the primary focus of both developers and implementers of CIS/2 has been on the pre-construction processes of design, analysis, detailing, and shop fabrication. This paper examines the applicability of CIS/2 to on-site construction processes, focusing on automating the erection and surveying of structural steelwork and integrating these two processes into the overall project delivery system.

KEYWORDS: automated construction, integration, product data model, project delivery, structural steelwork

Introduction

Successful integration of an automated process into a project delivery system depends on the ability to describe the desired result of the process in a manner understandable to its automation technology. Similarly, it depends on the ability of the automation technology to describe the actual result of the process in a manner understandable to subsequent processes. In this paper, the second release of the CIMsteel Integration Standards (CIS/2) [1] is examined for its ability to specify these descriptions in the specific cases of automating the erection of new structures and the surveying of existing structural steelwork.

The CIS/2 has been endorsed by the American Institute of Steel Construction (AISC) as the standard for the electronic exchange of structural steel project information for the North American structural steel design and construction industry [2]. The CIS/2 is derived from the deliverables and experiences of the Pan-European Eureka CIMsteel project and is intended to specify the exchange of data and the integration of engineering applications across steelwork design, analysis, and manufacturing.

A number of software vendors with structural steelwork applications are developing and demonstrating CIS/2-compliant translators for their products as part of the AISC Electronic Data Interchange initiative [3]. Initial case studies have shown significant benefits resulting from the use of these translators to exchange data between software applications in real projects. These benefits include knowing earlier in the project timeline approximately how much and what types of steel need to be reserved, reducing the time needed to estimate the material cost as part of the bidding process, and saving weeks of detailing time in a typical project [4].

The CIS/2 characterization of steelwork is intended to cover the needs of on-site construction as an aspect of manufacturing. However, the emphasis in the CIS/2 documentation is clearly on the pre-construction processes of design, analysis, detailing, and shop fabrication. None of the existing implementation efforts specifically addresses the information needs of field erection, although some do involve applications that can assist construction planning (for example, through the breakdown of the structure into erection zones and the identification by zone of the members to be shipped).

In order to automate the field erection of a structure as an integrated process in the overall project delivery system, one needs to access computer-sensible descriptions (1) of the parts and prefabricated assemblies that are to be delivered to the construction site, (2) of the sequence in which these parts and assemblies are to be assembled during erection to create the final structure, (3) of the joint systems connecting these pieces together, and (4) of the positions and orientations of all the pieces in their final locations. Automated construction equipment can then proceed to erect the structure without human direction.

Likewise, to automate the surveying of an existing structure as an integrated process, one needs a computer-sensible description of a model of the steel pieces composing the structure that can be populated with piece identification and position data acquired in the survey using various metrology systems.

In the following sections, the ability of the CIS/2 to provide these descriptions and thus to support the automation and integration of the erection and surveying of structural steelwork is explored.

Overview of CIS/2

The following is an extremely brief overview. The specification [1] should be consulted for details about what is said here, and for all the topics that aren't discussed here, such as the definitions of features and fastener mechanisms.

The CIS/2 was developed using some of the methodologies and technologies of the international product data standard ISO 10303 [5], known familiarly as STEP. The information requirements for relevant concepts in the domain of structural engineering were captured in structured English. After these information requirements were agreed, they were interpreted in a formal, computer-sensible model known as the Logical Product Model (which is LPM/5 in the second release of CIS). The information constructs in LPM/5 are defined as a schema using the ISO/STEP information modeling language known as EXPRESS [6]. This schema represents the constructs in terms of their underlying data entity types, attributes, relationships among entity types, and constraints.

An implementation of CIS/2 in a software application proceeds first by developing mappings between the concepts represented in the application and like concepts defined in the LPM/5. Depending on the particular concept(s) involved, a mapping may be as simple as one-to-one or as complicated as many-to-many with associated constraints. Translators for importing and exporting information to and from the application in the form of CIS/2 exchange files are developed based on these mappings and the ISO/STEP clear text encoding method [7]. This method establishes the syntax of the exchange file and the rules for composing exchange structures in that syntax. It should be noted that the exchange structures can not be interpreted without recourse to the LPM/5 which establishes their meanings.

The LPM/5 supports three principal views of structural steelwork information, represented as analysis models, design models, and manufacturing models. Data representing one, two, or all three views may be present in an exchange file.

Analysis models of the structure are built up from nodes and elements and support a number of different static and dynamic analysis methods. Further, analysis results may be given at nodes, element ends, or at a point within an element, and the results can be described in a variety of ways.

Design models represent the structure as a design assembly for the purpose of member and connection design. Design assemblies can be decomposed into other design assemblies and ultimately into design parts and design joint systems, which are the conceptual representations of a basic piece of steel and a basic joint system, respectively. Design results for members and connections can be described in terms of elastic or plastic resistance.

Manufacturing models represent the structure as manufacturing assemblies for the purpose of detailing, production planning, and manufacturing. Located assemblies are built up from located parts and located joint systems, which are the representations of a basic physical piece of steel and of a basic physical joint system (e.g., a fastener group or a weld), respectively. In turn, all these located items can be built up into larger located assemblies, culminating in the complete structure. Features also are located with respect to the item they modify.

The located parts and located joint systems and their higher order located assemblies represent the items that actually occur in the completed structure. The material and geometrical definitions of these items are separated into part, joint system, and manufacturing assembly definitions that are referenced by their respective located items. This separation of data facilitates the economical reuse of generally verbose definition data across many occurrences of what are often identical items.

Although the LPM/5 supports the explicit description of geometry using the ISO/STEP geometric resources [8], most CIS/2 parts are expressed using implicit descriptions of geometry based on the specification of a transverse section profile (or a sheet width) and a longitudinal length through which the section profile is swept.

The LPM/5 defines a hierarchical system of locations (e.g., positions and orientations) such that an item is located with respect to the higher order item to which it belongs. For example, a feature may be located with respect to a part, which in turn may be located with respect to an assembly, and so on up to the location of a structure with respect to a site and the location of the site with respect to the earth. Each location is defined using a specialization of the coord_system entity. For most CIS/2 manufacturing model instantiations, this specialization turns out to be the coord_system__cartesian_3d entity, which encapsulates the ISO/STEP geometrical entity axis2_placement_3d [8]. This entity has among its attributes a 3D point defining the origin of the coordinate system and a set of vectors defining the 3D orientation of the coordinate system.

The preceding information constructs allow the expression of information about structural steelwork in the form of data instances. They don't allow the expression of information about that data, e.g., of metadata related to its management over the life of a project. The Data Management Control (DMC) schema in the LPM/5 provides that capability. The DMC concepts are important to this study because they allow one to exchange and manage multiple versions of data instances. For example, they could be used to maintain simultaneously an "as designed" model of a structure, an "as built" model, and a current "as is" model, with the data instances interrelated in a manner similar to the content of a software change control and configuration management system. The details of the DMC are given in the CIS/2 [1]. For the purpose of this study, it is sufficient to know that the DMC schema allows the representation of such information as the unique identifier for a data item, the person who created the data item, whether the data item is new or has been modified, why it was modified, and the date of modification.

CIS/2 Applicability to Automated Erection

It should already be apparent from the preceding overview that the CIS/2 is applicable to the automated erection of new structures. Through the composition hierarchy of the manufacturing model, all the parts and prefabricated assemblies that are to be delivered to the construction site can be described, along with the joint systems connecting these pieces together. Through the nested coordinate systems that locate parts in assemblies and locate parts and assemblies into bigger assemblies, and so on up to locating a structure on a site, the positions and orientations of the pieces and all their features can be defined with respect to a geo-referenced global coordinate system. Of course, the global coordinate system used in the manufacturing model and the global coordinate system used in the automated erection process must be reconciled before proceeding with erection.

One complication that arises in using the LPM/5 manufacturing model is its extensive use of implicit geometry. A member may be represented by a manufacturing assembly composed of several plates welded to a wide-flange section with a number of clip angles on the section and on the plates. As stated in the overview, each of these parts---plate, section, and clip angle---is normally represented implicitly through the specification of a transverse section profile and a longitudinal length. Any datum of interest, sometimes called a principal point or a fiducial point, on the member must be calculated from these specifications and the (possibly) nested coordinate systems. However, this is really just a detail to be taken care of in software.

Another complication that arises is the need to be able to map between a piece physically present on the site and the representation of the same piece in the manufacturing model. A unique piece mark can serve to identify the two items but this is insufficient to ensure the items are oriented the same way in both worlds. An obvious way to ensure this orientation is to create a rule-based convention for placing the piece mark in a specific location on the physical item. This is similar to the rule-based conventions found in software that lays out members in a common way in 2D fabrication drawings, such that the left end and the front face of the member as shown in the drawing always bear the same relationship to the member.

Finally, the LPM/5 location mechanism is based on the STEP axis2_placement_3d entity, which encapsulates the 3D point of origin and the 3D orientation of a Cartesian coordinate system. With some automated placement techniques, this is just the right information. With others, the information needed may be, e.g., the set of 3D points defining all the corners of the member. Again, this is a detail to be taken care of in software.

CIS/2 Applicability to Automated Surveying

Two different situations arise in the surveying of existing structural steelwork. In the first, a CIS/2 representation of the structure in the form of a manufacturing model already exists. It might be the "as required" model provided by the project design team, for example. The member identification and location data measured in the survey can be used in this situation to update the existing CIS/2 representation to document the "as is" condition of the structure. Automation of this process would be relatively straightforward.

If the surveying application supports the DMC schema, then this "as is" representation could be achieved economically using DMC constructs to record the survey measurements with respect to their equivalent measures in the existing model. This process can be thought of as the electronic version of "red lining" a drawing. With DMC implemented, even fragmentary survey measurements can be associated properly with the model.

If the DMC schema is not supported in the survey application, then the existing CIS/2 representation would have to be duplicated and the copy updated to account for the survey measurements. Because it results in an entirely new model, this is a less desirable approach. It is a particularly undesirable approach if less than 100 percent of the structure is surveyed. In this case it becomes difficult to build a consistent model because the CIS/2 manufacturing model does not provide any versioning capability and hence can not make a distinction between original and updated location values.

Of course, the global (presumably geo-referenced) coordinate systems used in the existing manufacturing model and in the survey process must be reconciled before any model updating can occur. As well, the nested coordinate systems present in the manufacturing model must be taken into account properly. This nesting is dependent on the composition hierarchy used in creating the manufacturing model and that hierarchy may not be apparent to the survey process. Hence, the most transparent approach would appear to be to make the survey measurements in the global coordinate system and transform the location data for each survey datum in the manufacturing model to the same global coordinate system for comparison. As described in the previous section, the datum may have to be computed from an implicit geometrical description of the item in question. For each datum, the inverse of the transform then can be used to transform the relevant survey measurement back into the nested coordinate systems as required for updating the model. Finally, all survey data of parts and assemblies must be reduced to the LPM/5 location mechanism, based on the axis2_placement_3d entity, which encapsulates a 3D position and a 3D orientation as opposed to a collection of 3D positions of the end points of a member.

In the second situation arising in surveying, a CIS/2 representation of the structure does not already exist but must be created from the survey measurements. This task could be accomplished directly in the CIS/2 representation with a suitable new software application. Equally, the application could be an existing structural steelwork detailing package with a CIS/2 translator. Either way, the surveying application must build up a model of the parts and assemblies in the structure from detailed measurements of the members without *a priori* knowledge of the nature of the members, making automation of the surveying process difficult. In this situation, however, the composition hierarchy is known explicitly and the location data can be transformed appropriately.

General Issues

Establishing that the information structures needed to support construction site processes are present in the LPM/5 is a necessary but not sufficient condition for success. These information structures must be implemented in the CIS/2 translators for the attendant software applications. The CIS/2 use conformance classes to document implementation requirements. The conformance classes already developed in the CIS/2 need to be reviewed for adequate coverage of construction-related information requirements and then used to specify the required translator capabilities.

In common with most significant product data standards, the LPM/5 sometimes provides more than one way to specify the same information. Where found to be present, these redundancies must be circumscribed through implementers' agreements so that misinterpretation of exchanged information is minimized.

Also in common with most significant product data standards, the LPM/5 contains many weakly specified uses of character strings as identifiers, labels, and descriptions. Again, where these are found to be present in the information structures needed to support construction site processes, they should be strongly specified in implementers' agreements so that misinterpretation of exchanged information is minimized. As an example of such an agreement, albeit at the national level rather than the level of individual implementers, the AISC has updated its naming convention for structural steel products for use in electronic data interchange so that software implementers no longer have to decide for themselves how to construct character strings to designate standard sections.

Finally, the use of CIS/2 to integrate automated processes in the construction of structural steelwork will still fail despite attention to all of the above general issues if the users of the CIS/2-enabled software applications don't provide the necessary information. As an example, the CIS/2 allow for the use of geo-referenced coordinate systems, and the conformance class approach taken in CIS/2 can be used to ensure the translators involved properly map to and from the LPM/5 representation of this information, but none of this can force the user correctly to geo-reference the manufacturing model of a structure, and virtually no one does today using their favorite structural steelwork detailing system. A series of recommended practices will need to be developed to guide the end user to ensure success.

Summary

The second release of the CIMsteel Integration Standards has been endorsed by the American Institute of Steel Construction as the standard for the electronic exchange of steel project information for the North American structural steel design and construction industry. In this paper, the applicability of the CIS/2 to on-site construction processes has been assessed, focusing in particular on the erection of new structures and the surveying of existing structural steelwork. The manufacturing model view of CIS/2 has been found to be capable of supporting these processes. With the ongoing implementation effort in the AISC Electronic Data Interchange initiative and with the development of modest new applications, these processes can be automated and integrated into the overall project delivery system using CIS/2 as the baseline specification.

REFERENCES

1. Crowley, A.J., and Watson, A. S., CIMsteel Integration Standards Release 2, Vol.1-6, The Steel Construction Institute, Ascot UK, 2000. (see also the companion website http://www.cis2.org)

2. Press release, American Institute of Steel Construction, Chicago, IL, 2001 (reprinted in Foreword of CIS/2). (see also the website http://www.aisc.org)

3. Hamburg, S. E., and Holland, M. V., Leaping Ahead with EDI, Modern Steel Construction, Vol. 39, No. 1, pp. 42-48, American Institute of Steel Construction, Chicago, IL, February 1999. (see also http://www.aisc.org/edi.html)

4. Electronic Data Interchange Case Studies, The American Institute of Steel Construction, Chicago, IL, 2001. (electronic version available from http://www.aisc.org/edi.html)

5. ISO 10303-1, Industrial automation systems – Product data representation and exchange – Part 1: Overview and fundamental principles, ISO/IEC Geneva, Switzerland, 1994.

6. ISO 10303-11, Industrial automation systems – Product data representation and exchange – Part 11: Description methods: The EXPRESS language reference manual, ISO/IEC Geneva, Switzerland, 1994. (with Technical Corrigendum 1, 1999)

7. ISO 10303-21, Industrial automation systems – Product data representation and exchange – Part 21: Implementation methods: Clear text encoding of the exchange structure, ISO/IEC Geneva, Switzerland, 2002.

8. ISO 10303-42, Industrial automation systems – Product data representation and exchange – Part 21: Integrated generic resource: Geometric and topological representation, ISO/IEC Geneva, Switzerland, 2000. (with Technical Corrigendum 1, 2001).

BEYOND WEBCAM:
A SITE-WEB-SITE FOR BUILDING CONSTRUCTION

Suchart Nuntasunti
Graduate Research Assistant, snuntas@unity.ncsu.edu

Leonhard E. Bernold
Associate Professor
Director, Construction Automation and Robotics Laboratory, http://carl.ce.ncsu.edu
North Carolina State Univ., Raleigh, NC, 27606
bernold@eos.ncsu.edu

ABSTRACT: The use of project website and web cam technology in the construction industry is limited to providing basic information about the project and 24/7 pictures of the site. This paper discusses the promises of a site-based website that serves as the central hub for real-time communication, monitoring, and control in building construction. Such a website offers opportunities that go further beyond what is in use today. The paper lists a series of possible applications that have been tested at several construction sites in Raleigh, North Carolina. In addition to helping contractors plan and control the resource flow in real time the site-web-site is also used as a tool for creating daily reports, rapid problem solving, remote inspection, productivity measurement, on site security, communicate with on-site equipment (e.g., in emergencies), and to create visual as-builts.

KEYWORDS: collaboration; productivity; Internet; logistics; value; website; equipment; safety

Introduction

According to U.S. Census Bureau, in 1997 the construction industry performs more than $845 billion of work or almost 11% of GDP and employs more than 5.6 million employees. Thus, construction productivity is important for the growth of entire U.S. economy. However, statistics show that productivity of the construction industry has been far below other industries for the past three decades. In addition, productivity is not the only problem facing the industry. Quality of work, waste, and safety are among the issues that decline the industry's overall performance. Therefore, increasing productivity, improving quality, minimizing waste, and providing a safe environment for workers are main the goals in managing a construction project. In order to achieve these goals, attention must be paid to the planning and controlling of the processes rather than on assessing, measuring, and repairing after the product is finished. In other words, it is only before and during the actual operation that management has the chance to intervene and correct, ensuring optimal performance under the given circumstance. This, however, requires real-time and accurate information coming from different sources.

With today's digital and world-wide-web technologies, communication has never been easier and faster. The construction industry has adopted web technology and some projects already have their own websites. Nevertheless, the application has been limited to providing basic information such as project participants, construction area, and contact information. Some projects also have a web cam to provide real-time pictures of the site. However, by integrating the web with other technologies such as wireless, sensor, and tracking devices, the project web site can be transformed into a powerful tool for every participants to plan and control. Because the potential use of the technology is limitless, it is imperative to investigate the benefits they could provide to each user. This paper discusses the promises of a site-web-site (SWS) that serves as the central hub for real-time communication, monitoring, and control in building construction. Such a website offers opportunities that go far beyond the conventional project website and web cam.

Coalition Building: A Win-Win Strategy

The SWS is designed for six major users: 1.) Owner, 2.) Designers, 3.) Inspector, 4.) Contractors, 5.) Subcontractors, and 6.) Suppliers. Prerequisite of the successful SWS is open information sharing between participants. Unfortunately, the sharing of information is very uncommon in an industry that is constantly squeezed by the low-bid mentality of the owners. Low-bids foster a zero-sum culture where everybody fights for a piece of a small pie without considering how his piece is connected to the much larger pie. If we can't change the zero-sum strategy of contractors, the main benefit of web-based communication network will be lost. Thus, this paper will first develop a concept that could entice contractors, suppliers, and subcontractor to build coalitions as a framework to share information using a site-web-site (SWS).

Collaboration between contractors and suppliers of a building project offers to drastically improve productivity and cut resource waste. The basic premise of a contractor coalition is to cut the cost of operation for the participating coalition partners by optimizing the use of the combined resources. Figure 1 presents a simplified comparison between a competitive and collaborative approach. As indicated, the overall benefit or the value of collaboration is $20 ($150 - $130). This gain, however, can only be achieved if the supplier spends $10 more on his operation when compared to the traditional method. It is apparent that the only beneficiary of the coalition is the contractor who shows a net gain of $ 30. In order to make collaboration attractive to the supplier, the contractor has to be willing to share the savings, which he really can only achieve with the added work of the supplier. If we assume that the two coalition partners are splitting the gain equally, the supplier will be reimbursed his $10 extra cost and receive an additional $5 from the "coalition fund" reducing his actual cost to $45 compared to the $50 in the competitive mode. The contractor, on the other hand, will end up with a cost of $95 vs. $100. Figure 2 highlights a real world application of this concept.

In a comparative field test Salim and Bernold (1994) showed that through up-stream planning, rebar could be bundled and shipped according to the way the rebar is placed. If the rebar is staged properly, as indicated in Figure 2, ineffective time spent on re-handling and searching can be eliminated. The result of the study showed that placement-oriented delivery and staging improved crew level productivity in the placement of rebar by 30% compared to the traditional method. Key to the productivity gain in laying the bars, however, was a supplier who was willing to collaborate with the subcontractor in shipping the steel in a way that matched how the crew was progressing.

Framework of the Site-Web-Site (SWS)

The main objective of the SWS is to providing an electronic hub that allows the real-time sharing of information between collaborating entities. The system consists of a notebook computer with broadband Internet access as a hub and a set of peripherals such as docking station for PDA, motion detector, and video camera. The camera used can be wire or wireless. The wireless camera provides more flexibility but the drawback is poor clarity of the picture. In addition, the series of digital picture have also been taken and keep in the website as a project history. The information maintained in the system is visual and media rich. The system has been tested at the residential building construction site in Raleigh, North Carolina. Furthermore, the extension of the system is limitless. As shown in Figure 3, the list of integrated transmitters, receivers, sensors, data entry ports, access portals, etc. can be extensive. It is also indicated that on-site data communication using wireless technology will make the system extremely flexible.

Applications of the Site-Web-Site (SWS)

The following section describes several possible applications for SWS.

Synchronizing the Resource Supply

The Council of Logistics Management (CLM), defines logistics as a part of the supply chain process that plans, implements, and controls the efficient, effective flow and storage of goods, services, and related information from the point of origin to the point of consumption in order to meet customers' requirements (CLM, 2001).

In addition to the physical material, information about the material such as: 1) bill of material, 2) purchasing order, 3) specification, 4) delivery date, and 5) staging area can also be modeled as a flow. Figure 4 presents a schematic that indicates that besides material, all the resources necessary to complete a process can be incorporated into a flow model that links the Point of Origin (POO) of a resource to a point of consumption (POC). It is apparent that the goal of a synchronized resource logistics model is to ensure the availability of all resources, at the right time, at the right quality, at the right quantity, at the right place, for the right price. As the example of placing rebar indicated, optimal resource logistics in construction offers many opportunities for collaborative cost savings, since many of the resources are controlled by independent business entities. Web-based resource logistics offers a unique technology that supports such coalition building.

24/7 Visual Site Access

An example website that includes the picture from a web cam is presented in Fig. 5. The interval time for refreshing or storing pictures can be set from 1 to 10 min. Some of the most common benefits of such a technology include:

- Real-time review of project status from anywhere w/internet access
- Automatic recording of environmental conditions, project progress, etc.
- Rapid problem-solving if visual information is needed
- Automatic surveillance at night if connected to a motion detector and light source

Visual Inspection

Many inspections required during construction are visual. Figure 6 features two digital images during and after the completion of the foundations for a single family home. In the same vein, SWS would allow company safety inspectors to do their visual inspection remotely.

Visual As-Builts

One unique opportunity provided by SWS is the creation of visually or electronically generated as-builts. Spatial data about the exact location of a new water pipe can be collected from a laser positioning system. Alternatively, digital pictures allow the new homeowner to see through the walls. Figure 7 exhibits, two examples how digital images can be used for establishing as-builts. If the inspection of the electrical and pluming system would include taking a series of digital pictures, an extremely useful information bank could be established. Another example of useful visual as-builts is the pictorial marking of buried utilities such as water, sewer, cables, and gas. Any homeowner who has changed the landscape around his/her house could benefit from the availability of such information.

Automatic Resource Tracking

The RFID (Radio Frequency IDentification) tag technology has successfully been used in the retail and service industries. For example, Wal-Mart and FedEx have implemented RFID tags to improve their supply-chain and logistics management. By using RFID tags, the SWS will be able to track and identify materials, equipment, tools, and other resources automatically.

As was discussed earlier, synchronizing the resource logistics must be one of the main goals of a concept that wants to create win-win partnerships. Since equipment, material and labor are key resources on a building site, tracking their whereabouts is essential. Figure 8 displays two situations where tracking is able to provide valuable information for different project participants. Figure 8 a) is of value to the framing and general contractor, as well as to the lumber supplier. For example, both contractors can estimate precisely when the framing crew can start and when the materials should be delivered. It would allow the framing contractor to indicate where he would like the lumber to be staged. He could do this by marking the digital picture and submitting it to the supplier. The effect would be similar to the rebar example discussed earlier in that unproductive transportation time could be eliminated.

Figure 8 b) shows that the trusses have been delivered but not necessarily in an optimal position. First, it will be necessary to separate the stacks of trusses. If no crane is available, this work has to be done by hand one truss at a time. Secondly, the staging location is not efficient because the distance to the footprint of the house is quite large. Most importantly, however, the framing contractor can see clearly the status of the framing work and the exact location of the delivered roof trusses.

BlackBox Technology

Electronic monitoring of large systems can have many different goals. One example is to detect an unsafe status of equipment while in operation. Bernold (at el., 1997) developed such a system for cranes. The intelligent monitoring system can be used to retrofit existing crane hardware. The simple architecture and the transportability of the sensors provide opportunities for utilizing the concept for other types of cranes or even for other machinery where unsafe conditions cannot be detected easily by an operator. Figure 9 depicts how a BlackBox mounted on a crane can be equipped with a wire-less communication interface to alert people or establish communication between equipment that are linked to the SWS system automatically. In addition, such information can be maintained in the server computer and be used as a historical record for the equipment.

Summary

Resource logistics, which was once limited to storing and moving of goods, has become a critical component of company management. Storing and transportation are still necessary but now these processes are seen as source of "waste" that should be minimized. In addition, logistics of goods alone is no longer sufficient. Information, considered itself a resource, needs to be part of logistics management. Because of the highly fragmented nature of the construction industry information is being created and needed by many different companies. The main obstacle to sharing information freely, however, is the low-bid orientation of the industry. Using a win-win coalition building strategy will create the necessary incentive to make a site-web-site an economically successful concept.

This paper presents a prototype website for building construction designed to help owner, contractors, suppliers, and other participants to plan and control resources in real time using 24/7 visual accesses to the site. Using a set of examples different uses and benefits of SWS were explained. The system has been tested at various construction sites in Raleigh, North Carolina. The results show promise to create an effective tool to improve communication between contractors, supplies, and homeowners. In addition to real time accesses, the system is also used as a tool for rapid problem solving, and increase security. The digital images taken at the site provide a historical record and "visual as-built" for the homeowner to show the location of utility lines both behind (inside) the wall and buried under ground.

Acknowledgements

The work presented in this paper has been funded by grant CMS 0080073 from the National Science Foundation. Their support is gratefully acknowledged. Any opinions, finding, conclusions, or recommendations expressed in this study are those of the writers and are not necessarily reflect the views of the National Science Foundation.

REFERENCES

Bernold, L.E., Lorenc, S.J., and Luces, E. (1997). "Intelligent Technology for Truck Crane Accident Prevention." J. Construction Engineering and Management., ASCE, Vol. 123, No. 3, September, pp. 276-284.

Council of Logistics Management (CLM). (2001). http://www.clm1.org

Koskela, L. (2000). "An Exploration Towards a Production Theory and Its Application to Construction." Ph.D. Dissertation, Helsinki University of Technology, Finland.

Salim Md. and Bernold, L.E. (1994). "Effects of Design-Integrated Process Planning on Productivity in Rebar Placement." ASCE, J. Construction Engineering and Management, Vol. 120, No. 4, December, pp. 720-739.

Schonsleben, P. (2000). "Integral Logistics Management: Planning & Control of Comprehensive Business Process" New York: St. Lucie Press / APICS.

Mode of Relation	Cost to Supplier	Cost to Contractor	Total Cost	Cost Saving
Competitive	$ 50	$ 100	$ 150	$ 20
Collaborative	$ 60	$ 70	$ 130	

Fig 1. Cost-benefit analysis of collaboration

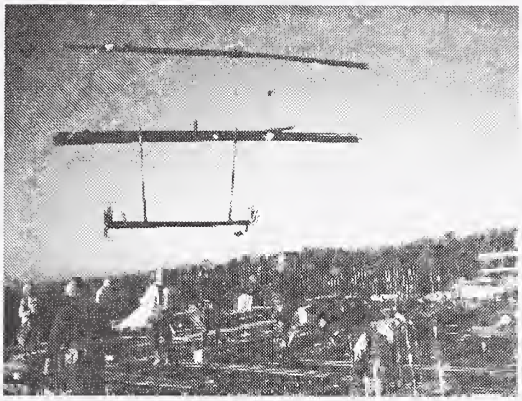


Fig 2. Rebar delivered according to the sequence of placement

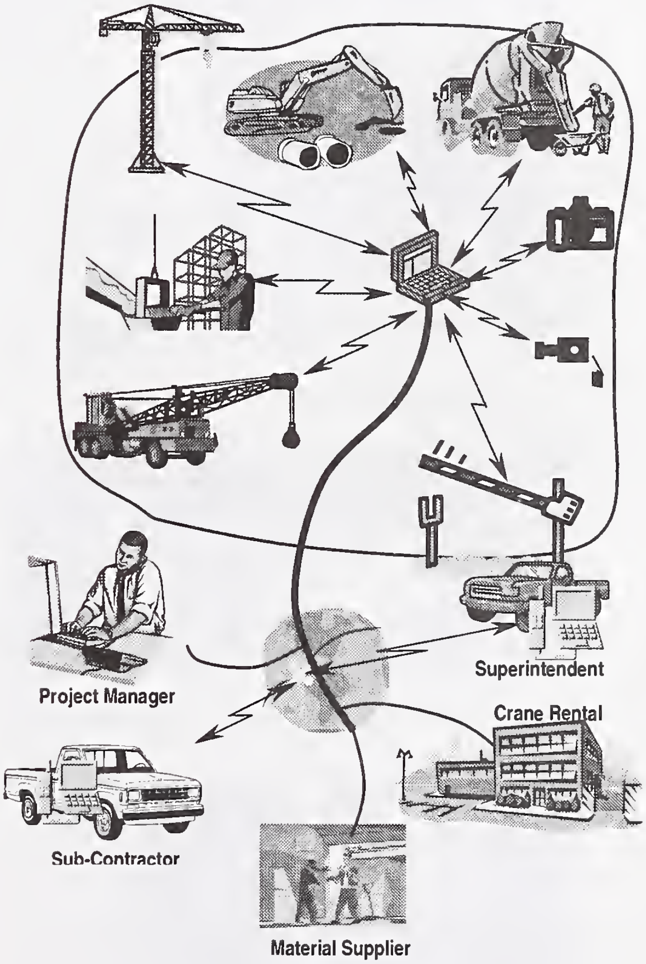


Fig 3. Communication Network of a Site-Based Website

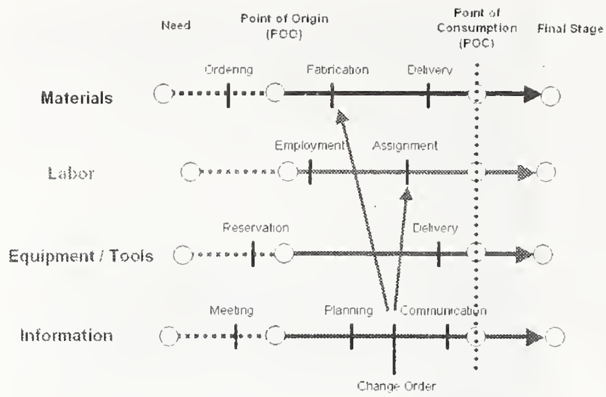


Fig 4. Integration of Resource Logistics

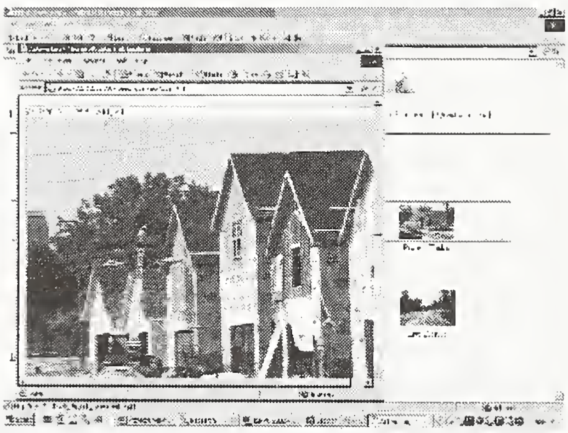
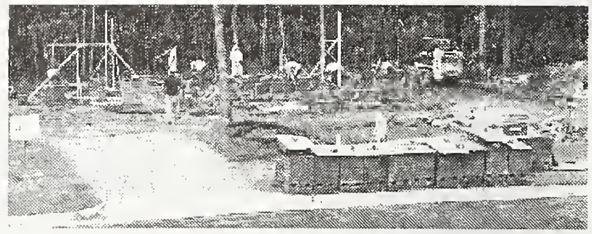


Fig 5. WebCam On a Site-Web-Site

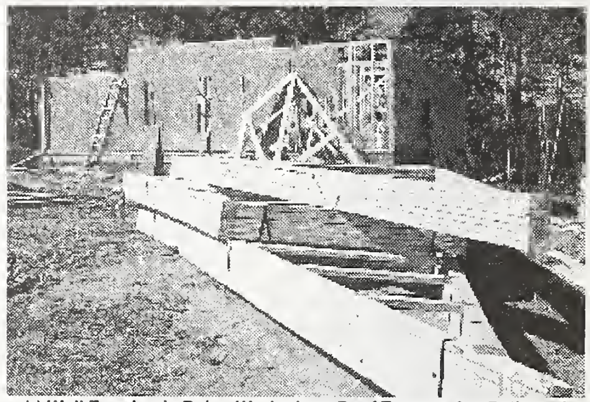
Fig 6. Remote Inspection of Foundation Using SWS

a) Studs, Pluming, Power b) Buried Water Line

Fig 7. Digital Pictures as As-Builts



a) Labor is Establishing Foundation Walls, Brick is Staged at Road, Bobcat is Operating



b) Wall Framing is Being Worked on, Roof Trusses Are Delivered

Fig 8. Tracking Status and Location of Resources

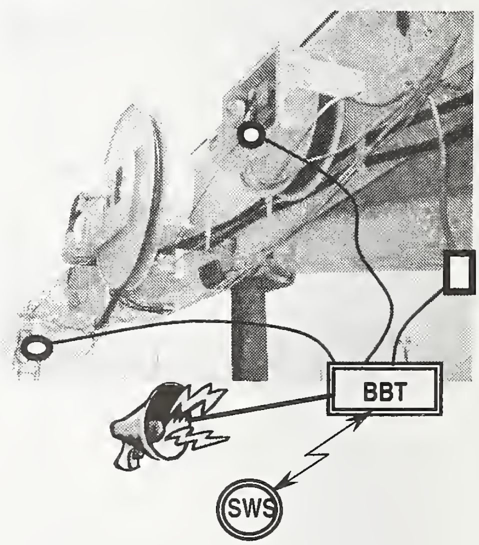


Fig 9. BlackBox Linked to SWS

DEVELOPING A CONSTRUCTION INTEGRATED MANAGEMENT SYSTEM

Yan-Chyuan Shiau, Ming-Teh Wang, Tsung-Pin Tsai, Wen-Chian Wang

Associate Professor
Department of Construction Engineering
Chung Hua University
Email: ycshiau@ms22.hinet.net

Abstract: Building market has been shortened for last few years and tremendously affected the construction profit in Taiwan. In order to survive in such abominable environment, people have put many efforts on construction methodology, materials, budget planning, and strategy alliance. However, to properly enhance the construction management is the key issue to raise company's competition. Therefore, an effective administrating tool is a must in construction life cycle such as estimating, bidding, procurement, scheduling, site managing, and valuation. This research has utilized Object Oriented concept, E/R techniques and Delphi environment to develop the Construction Integrated Manage System (CIMS). Through the help of this tool, user can effectively generate Bidding List, Purchasing Materials, Labor and Material Analyze Items and Specifications. Basic Database System can help us to modify resource prices to adjust budget, effectively control project cost and financial planning. Budget Control System will inform engineers when overspend occurs. Purchasing System can assist us to select the best subcontractors with most reasoning price. Valuation system can extract money request data from site records; therefore, prevent the over-calculating and short-calculating to achieve a better financial administration. With the help of CIMS, we can effectively control all details such as bidding, purchasing, contracting and evaluating, for each construction stage. This integrated system can practically control all information, maximize construction management efforts, raise company's strength, increase competition, and create profit for construction industry.

Keywords: Construction Management, Cost Control, Delphi, Object Oriented.

I. Introduction

With the rapid development of information technology, the traditional business behavior is deeply affected by the e-Commerce, and the management strategies of traditional industries are undergoing a fierce evolutional change. The features of the e-Commerce include:

A. Fast Speed – Real-time service, quick response and update.
B. Large Quantity – Simultaneity, automation and diversity.
C. Lower Failure Rate – Digital data transmission reducing personal input errors.
D. Consecutive Service – 24-hour and 7-day service for the whole year.

With the aforementioned advantages, the services rendered by e-Commerce are better in quality than the services rendered by traditional providers. The traditional service providers are much weaker in competitive strength in comparison with e-Commerce companies as they are restricted by manpower. Thus, it is required to respond to the development of this tendency to enhance the productivity and competitive strength of the traditional building industry

To assist the industry in achieving the aforementioned goal, the "Integrated Information Management System for the Building Industry" is developed to integrate vertically the estimate, price assessment, bidding, price comparison, construction management, price calculation, payment request, and cash flow management. All the relevant information during the life cycle of a project I, and the calculation result of each pre-operation is stored in the system for future use to avoid any errors caused by repeated calculations and inputs.

II. Structure and Function of the System

The "Construction Integrated Management System" comprises the following modules:

A. Basic Database
B. Budget Module

C. Procurement Module
D. Price calculation Module
E. Settlement Management Module

To integrate the sub-modules successfully, the study uses the object-oriented and network development tools to develop the modules. The function of each sub-module is described as follows:

2.1 Basic Database

Each building project is unique in its nature. Two buildings may have the same appearances, but the underlying bearing capacity and the conditions of neighboring houses are different from buildings to buildings. However, common characteristics exist among the components (column, wall, beam and formwork), materials (such as steel bar, concrete, ceramic tile), laborers (laborers in charge of reinforced steel, painting, plumbing) and documents (contract, construction instructions) used for a building.

Since such information is common for different projects, it is very useful to arrange it for future use so as to significantly educe the workload for data management and apply the cumulative experience for future projects.

The data contained in the database of the system include the basic resource data, work data, material analysis data, project instance data, subcontractor's data, labor data, project data, owner's data and database maintenance data (Fig. 1).

2.1.1 Basic Resource Database

The Basic Resource Database contains the data of manpower, machinery, materials, and other data concerning the resources to be used for the work.

2.1.2 Project Items

The project items are the data contained in bid documentations. In order to cope with complex information for a project, we need to use Object Oriented concept to mimic real world data by different objects. Based on the WBS [1], we make out the construction project items into three levels (Fig. 2): Large Items (15 items), Medium Items (341 items) and Small Items (3583 items). In this structure, Large Items are used for the high level items of the project (such as Reinforced Concrete Structure Engineering); Medium Items are used for the middle level items of the project (such as Systematic Formwork) and Small Items are used for the low level tasks of the project (such as DOKA column formwork).

2.1.3 Material Analysis Data

The cost of a project is acquired by analyzing the costs of material, labor and machinery. All the quantities and types of material, labor and machinery for are created in the system for the use of decision makers.

2.1.4 Project Data

In addition to the numbering, name, location and type of a project, the project data contains the data about the client number, contract number, commencement data, finish data, timeframe, contract price, contractor number, and designer name.

2.1.5 Subcontractor's Data

The Subcontractor's Data contains the data about (Fig. 3):

A. Basic Data (such as name, address, uniform number, capital and performance)
B. Business Items (Set upon the selection from the database)
C. Evaluation Data (including construction capability, reputation and financial status)
D. Unit price (Quoted by each subcontractor)

2.1.6 Owner's Data

The Owner's Data provides the data about the unit number, telephone number, responsible person, address, e-mail address, uniform number of the construction company concerned, project name, contact person, remark and other data related to the owner.

2.1.7 Database Maintenance Data

The cost structure differs significantly from projects to projects due to the difference in the location, timeframe, working environment, and nature of projects. To reflect the actual status, the user is allowed to set different construction conditions for different projects. The variables are adjustable according to the budget of each project and, in doing so, the user may control the progress of the construction. When the unit price of any resource changes, the user may adjust the cost and budget of the project concerned easily with the assistance of the internal calculation of the database.

2.2 Budget Module

The Budget Module is used to calculate the construction cost based on drawings, rules and other documentations. The system uses individual works of the project concerned as a basis to estimate the quantities of labor, machinery and material required for each unit, which will further be used as a basis for the procurement and price calculation operations.

The Budget Module is designed based on an open structure to allow the user to define the smallest budget unit in accordance with awarding and procurement. The awarding can be conducted by combining the labor and material, depending on the nature and works involved. The Budget Module of the system provides the functions of budget preparation, budget adjustment and bid document preparation.

2.2.1 Budget Preparation

The system divides the budget of a single work into several categories to reflect the actual status. The divided budget includes the construction budget, awarding budget, execution budget and closing budget. The system creates different project budgets based on selected databases, work instances and reversed designs.

2.2.2 Budget Adjustment

In general, the budget may be affected by some factors appearing during the budget preparation phase and needs to be adjusted. The system allows the user to adjust the budget by one of the following approaches:

A. Adjust the total work price based on a fixed percentage.
B. Estimate and evaluate the possible change in price index in accordance with the time set in the Construction Progress Sheet for different resources.
C. Adjust the unit price of a resource that may be affected by external elements and cause the price to rise significantly. The overall change of the work can also be calculated in this step.

2.2.3 Bid Document Printing

The Bid Document Printing function of the Budget Module is used to print bid documents. All documents of a project, or the data of individual works, units, houses or buildings can be printed separately or integrally without problems.

2.3 Procurement Module

The procurement is the first step for a company to gain profits. The management may use the Procurement Module provided by the system to control dynamic market prices, conduct price negotiation, build advantageous niches and create the highest profits. The Procurement Module provides the functions of:

2.3.1 Price Enquiry

Before the awarding is conducted, it is required to analyze the instructions, quantities and rules of the project and print relevant data, so as to control the site conditions and other information before quotation is made and minimize the possible problems that may be encountered during the construction. The system is capable to integrate the resources required for the project and analyze relevant works and materials before sending the price enquiry sheet. It will also prepare construction instructions and unit price analysis sheets for subcontractors to arrange their resources. The system provides the following two price enquiry approaches:

A. Announcement via Network: The system will announce the awarding information via network, including the name, location, works, scale, rules and instructions of the project. The subcontractors may register, enquire and print on line.
B. Price Enquiry Sheet Printing Operation: The price enquiry sheet provides the information of awarding, project type, quantity, and unit price analysis and construction instructions. The printed price

enquiry sheets will distributed to subcontractors so that they may arrange their resources accordingly (Fig. 4).

2.3.2 Price Negotiation and Comparison

The quotation of the subcontractors input in the system and the data imported by the subcontractors on line are integrated to make a price comparison sheet. After the identities of the subcontractors are confirmed, the price comparison sheet and the awarding budget will be considered carefully to form a basis for price negotiation.

The system will integrate the data provided by subcontractors and calculate the total score of each subcontractor objectively based on different evaluation criteria and weights (including the construction capability, management capability, financial status, reputation and market advantages) [2].

2.3.3 Contracting Documentations

This function is used to investigate what projects and works a subcontractor has ever performed and the experience, financial status and management capability of the subcontractor.

2.3.4 Material, Labor and Machinery Requirement Sheet

This function is used to print the type, specifications and quantity of the material, labor and machinery required by a subcontractor to perform the work.

2.4 Price calculation Module

The Price calculation Module is used to transfer the data about the works and quantities from the site daily report system, and perform the price calculation for the current works and quantities of the site. The system will read the data of the contract quantity, procurement counts and cumulative payment request counts for cross comparison and analysis. When finding that the payment request counts exceed the procurement counts, the system will give an alert. The management may control the expected and actual construction status from the system and prevent the occurrence of possible problems as early as possible to minimize the risk of loss.

III. System Development

The system is developed through the following steps:

3.1 Analysis of Existing Systems and Collection of Relevant Information

The existing construction management systems available in the domestic market are investigated in the study to find out their shortcomings and the solutions to meet the demands of construction companies. The study also analyzes the historical records of different projects to investigate the difference in their nature and create an integrated system that is capable to handle multiple projects.

3.2 Analyses of System Features and User Demand

The process analysis technique is used to determine the functions and information flow processes required for the system. All unnecessary functions or processes are removed to simplify the maintenance of the system. The safety system is analyzed thoroughly by investigating individual potential problems and risks of safety, their affection and possible solutions.

Surveys are done to users about their demand for the "Integrated Information Management System for the Building Industry" to determine the requirements for each module, data type, printing format and other functions, in the hope to provide the best interface operation features for the industry.

3.3 Creation of the Basic Data Framework

The data processed by the system comes from the database. Therefore, it is crucial to create an effective integrated data framework. In addition to referring to the database structure of existing construction management systems, the study also takes the habits of users into consideration to create a basic data framework that are appropriately integrated for each module.

3.4 Programming of Each Subsystem

The study uses Delphi and BDE Administrator as the development tools and makes use of the network database writing technique to write subsystem programs. In addition to the single PC operation mode and the client/server design concept, the system integrates the Internet and multi-platform features. The standard Windows™ interface format is used to minimize the difficulty in operation.

3.5 Test of System Functions

In addition to confirming that the functions designed for each module can run as expected, the functions are tested to make sure that the system can meet the demand of the industry. The practical instances of the industry are collected for the test, and the comments of the users are analyzed for the modification of each subsystem. The integrated test is performed after all the functions of each subsystem have been tested and confirmed. The objective of the integrated test is to make sure that the data stream runs smoothly between different modules. The users confirm it after trial use that the system is capable to demonstrate the functions set for the industry and meet its demand.

IV. Database Structure

4.1 Classification of Works and Analysis of Materials

The work data in the database is divided into three levels: Large Items (15 items), Medium Items (341 items) and Small Items (3583 items). Users may use the work data to create their own project data easily. The system also provides the material analysis sheet for each work, and all the labor, machinery and material fields in the material analysis sheet have their own resource numbering and quantity data. The users may change or update the data stated in the material analysis sheet from time to time and aggregate them to acquire the unit price of the work concerned.

4.2 Engineering Resources

The system contains 1100 resources and is capable to analyze the resources required for each work. The users may use the system to improve their finance management capability effectively, control the budget before commencing the procurement procedure, and prevent the cost from exceeding its allowed range. The database structure of the system is built with ER/Studio (Fig. 5).

V. Conclusion

The system is capable to integrate the tasks of a project, including planning, design, estimate, budget, procurement, documentation and price calculation. It is a very useful tool for the management of a construction company to control its projects effectively. The system provides the following functions:

5.1 Basic Database

The system divides the works to 17 large items, 341 medium items and 3583 small items, and integrates 1200 types of resources.

5.2 Budget Module

The system creates the project budget, awarding budget, execution budget and settlement budget based on the data in the database, individual projects and the latest plans. It also provides three approaches to adjust budget plans.

5.3 Procurement Module

The system provides a very useful technique for the selection of subcontractors and the determination of their weights, and allows users to announce the project to be awarded via Internet. The users may also use the system to print the quotation submitted by subcontractors for price comparison and negotiation, and select appropriate subcontractors in a fair and objective manner.

5.4 Price Calculation Module

The Price Calculation Module provides the works and quantities based on the data in the daily report system, and performs the price calculation for the current works and quantities of the site. The system will read the data of the contract quantity, procurement counts and cumulative payment request counts for cross comparison and analysis.

With the assistance of the system, construction companies may effectively control the information they need and integrate vertically the budget development, procurement and price calculation to carry out the construction management more efficiently. The system is capable to integrate the information scattered at different sites via Internet in a real-time manner to optimize the construction management of the construction company concerned, improve its competitive strength, enhance its organization, and bring it more profits.

VI. Acknowledgement

I would like to take this opportunity to show my appreciation to National Science Council, Executive Yuan, for its support of the study. Project No.: NSC 90-2211-E-216- 015

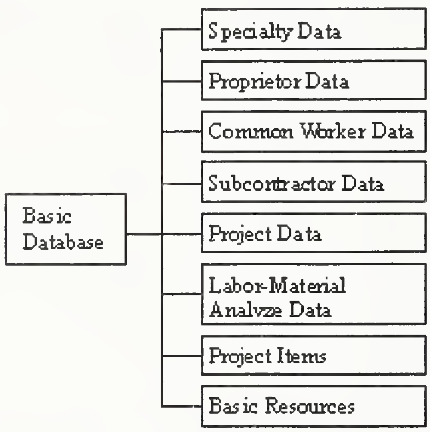


Fig. 1 Data Structure of the Basic Database

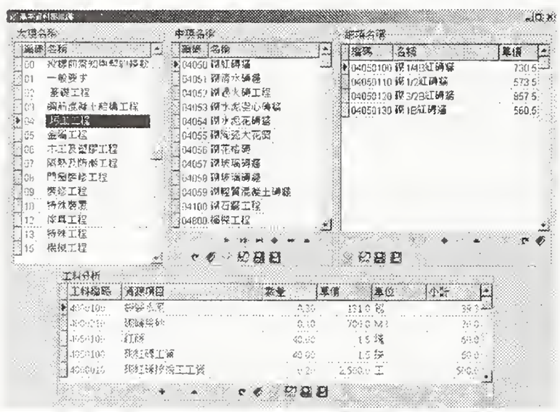


Fig. 2 Construction Project Dialog

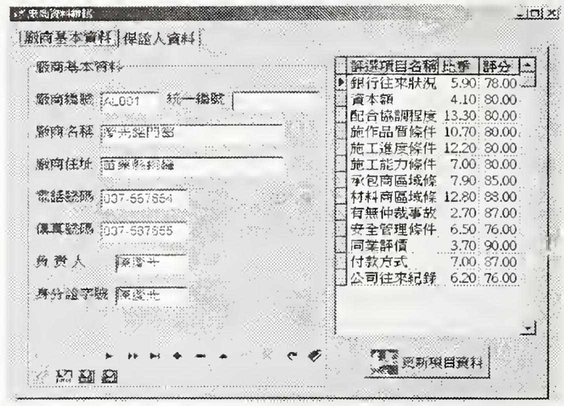


Fig. 3 Subcontractor Maintain Dialog

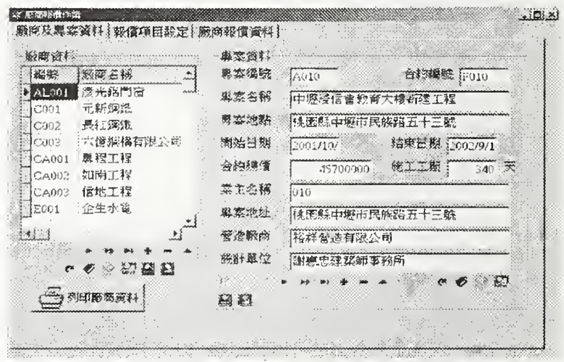


Fig. 4 Subcontractor's Quotation Dialog

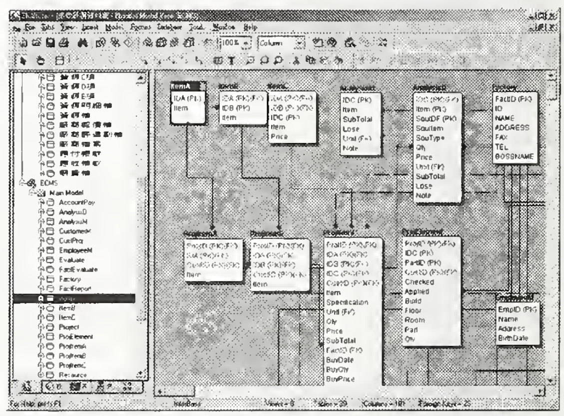


Fig. 5 Database Structure of the System

VII. REFERENCE

1. Xing-Dao Zhang, "WBS, the First Step to Integrate Cost and Schedule", Construction Management, pp40~44.

2. Zong-Bin Cai, "Research on Procurement for Building Projects", A Master Dissertation, Department of Construction Engineering, Chung Hua University, Xin-Zhu, Taiwan,

pp55~56, 2001.

3. Fouad M. Al-Sinan and Donn E. Hancher"Facility Project Delivery Selection Model,"Journal of Management in Engineering, ASCE, Vol.4, No.3, pp224-259, 1988.

4. Jian-Yu Chen and Zhong-Cheng Luo, "Contract Management for Government Project", Universal Law-108, pp29~30,1999.

5. Xian-Zhang Wu, "A General Discussion for Construction Contract Problem", Construction World, pp37~41,1999.

USE QUESTIONNAIRE AND AHP TECHNIQUES TO DEVELOP SUBCONTRACTOR SELECTION SYSTEM

Yan-Chyuan Shiau, Tsung-Pin Tsai, Wen-Chian Wang, Miao-Ling Huang

Associate Professor
Department of Construction Engineering
Chung Hua University
Email: ycshiau@ms22.hinet.net

Abstract: The construction scale is growing rapidly; therefore, to subcontract is a common style in Taiwan's construction industry. Selecting appropriate subcontractors is a key to assure the success of a construction project. With different environment, construction project usually invoke complex attributes. We need a good tool to help engineers to pick the best selections during various and complicate market. In this research, we use questionnaire to survey about 400 construction companies, which help us to obtain subcontractor selection factors and their weights. Subcontractors can fill available price for specific items from Quotation System, meanwhile, Appraisal System will integrate and calculate related data to determine total grades for all factories. Subcontractor Selection System can be used to compile all information about special area. Through the help of this system, Construction Company can use Internet to announce procurement, collect factory quotation and perform subcontractor selection. This can make the procurement process be fair, pick better partners, increase competition and create profit for company.

Keywords: AHP, Bid, Procurement, Quotation, Subcontractor Selection System.

I. Introduction

1.1 Motive of the Study

Comparing to other phases, such as planning, design, construction and so on, procurement consumes shorter time in the life cycle of a construction work. However, it is the key phase in which the cost of the project is determined. The selection of subcontractors often encounters problems, such as the selection of inappropriate subcontractors, difficulty in the management of subcontractors and out-of-control of budget and quotation systems. Such problems might be caused by insufficient time for execution, complicated procedures or poor information channels. It is, therefore, important for construction companies to control the subcontractor selection operation and make sure to conduct it in a fair and objective manner.

1.2 Objective of the Study

The objective of the study is to analyze the aforementioned problems so as to:

A. Investigate the status quo of the subcontractor selection system in Taiwan;

B. Acquire the evaluation criteria and their weights by doing surveys and via AHP; and

C. Develop a Subcontractor Selection Management Aid System, including Basic Database, Budget Management Module and Subcontractor Selection Module.

1.3 Scope of the Study

The study is limited to the procurement and subcontractor selection operations of a construction company. Construction projects are used as the subjects of the study.

1.4 Methodology and Process of the Study

The method and process used in the study is shown in Fig. 1.

II. Study of the Subcontractor Selection

There are some problems in traditional subcontractor selection strategies:

A. Engineers always pick familiar subcontractors, therefore, cannot get the best bargain for company.

B. The purchasing message can only reach to limited subcontractors.

C. People can easily collude with subcontractor and commit cheat in close environment.

The subcontractor selection operation includes the determination of evaluation criteria and their weights. This is the most difficult operation to conduct effectively and fairly in the procurement phase.

The importance of subcontractor examination is described in [1]. The study finds that it is important to establish methodologies for the examination system and design a simple and easy-use examination sheet for the of subcontractor selection.

Huang, Chung-Fa divides the subcontractor management into six subtopics in [2] and makes comparison among them. The six subtopics are registration of new subcontractors, final selection of subcontractors, examination of subcontractors, payment, handle and remove of disputes, and scale economy of procurement. The selection of subcontractors is described as follows:

2.1 Subcontractor Evaluation Form

The subcontractor evaluation comprises "Evaluation Facets" and "Coordination of Facets and Evaluation Principles". In [2], the evaluation criteria and their weights are divided based on Work Quality, Progress Control, Cooperativeness, Safety Management, and Material Management. Some construction companies divide the evaluation criteria into 12 facets.

2.2 Evaluation and Scoring

Subjectivity is the factor unlikely to be eliminated in the evaluation of subcontractors. Different evaluators may evaluate one subcontractor differently. To prevent the objectivity of the evaluation from being affected by subject factors, it is required to establish a reliable scoring system for the evaluation.

III. Evaluation Criteria and Weighing

3.1 Evaluation Criteria and Their Weights

To acquire an objective analysis result, surveying is used to investigate the evaluation criteria and their weights. The perfect score is 100. A fair, open and appropriate selection model is established by using the computer aided automatic calculation function in conjunction with arranged relative weights. It is used as a tool to select appropriate subcontractors in the awarding phase.

The flowchart of the evaluation criteria and their weights is shown in Fig. 2:

3.2 Application of AHP

3.2.1 Introduction of AHP

Analytic Hierarchy Process (AHP) is a decision analysis approach developed by Thomas L. Satty in 1971 [3]. As an easy-use and very practical tool based on a simple theory, AHP is capable to extract the comments of multiple experts and decision makers, and is mainly applicable to handling the problems arising in an uncertain environment in which multiple evaluation criteria exist. The AHP are used to systemize complicated problems and dissolve these factors into different levels from various directions. A comprehensive analysis is conducted via the process of quantification to assist decision makers in the selection of appropriate plans.

Many factors must be taken into consideration when operating the evaluation mechanism. The study uses AHP to establish a hierarchical structure for all affecting factors and acquires the weight of each factor by pair-wise comparison. The acquired weight distribution is more objective than setting the weight for individual factors.

3.2.2 Steps of AHP Analysis

The study uses AHP to conduct the decision analysis with reference to [4] and [5]. The steps of the analysis are described as follows:

The study uses AHP to conduct the decision analysis with reference to [4] and [5]. The steps of the analysis are described as follows:

1. Establishment of the Hierarchical Structure

   After the final goal of establishing the hierarchical framework is achieved, a mutually independent hierarchical relationship is built by interviewing experts and doing surveys (secondary goal), and analyzing the elements that might affect the secondary goal. The elements of similar importance are collected on the same level in this step. (Fig. 3)

2. Weight between the Elements on Different Levels

The calculation of the weight between the elements on different levels is completed though the following four steps:

A. Establishment of Pair-wise Comparison Matrix:

   The element comparison is conducted in this step. The parent element of an element on a lower level is used as an evaluation criterion for the pair-wise comparison.

B. Calculation of Priority Vector

   Divided each comparison value by the sum of the values in corresponding fields for the aggregation of the rows; namely, the sum of the percentage each comparison value occupies in its corresponding row.

$$\sum_{i=1}^{n} \frac{cell\_value_i}{column\_sum_i} \qquad \text{..................(1)}$$

   Formula (1) shows the sum of the percentage each comparison value occupies in its corresponding row. An n x 1 matrix is acquired in this step.

C. Calculation of the Maximum Eigenvalue .max:

   Multiply the entire matrix with the acquired priority vector to produce an n x 1 matrix. Then divide this matrix by the priority vector to acquire unit vectors. Calculate sequentially the average of the unit vectors to acquire the maximum eigenvalue .max.

D. Examination of Consistency:

   During the pair-wise comparison, discrepancies might occur between the results of the comparison and the decision. The consistence ratio of Satty's AHP is used to examine the consistency of the entire matrix.

3.3 Arrangement of Evaluation criteria and Their Weights

Surveys are done to 400 contractors and the consistency of the weight analyses is examined in the study. When the consistency is confirmed, the arrangement of the "Relative Weight Analysis of the Secondary Goal" and the "Relative Weight Analysis of the Evaluation Criteria" is conducted to acquire the overall relative weight. Table 1 shows the relative weight of the entire evaluation criteria.

IV. Development of the Subcontractors Selection System

The system mainly comprises the maintenance of subcontractor's data, quotation module, subcontractor selection module, budget control module, report preparation and maintenance of relevant data.

4.1 Maintenance of Subcontractor's Data

The "Maintenance of subcontractor's data" provides the basic data of the subcontractor and the data of guarantors. The user is requested to input the data of the subcontractor (Fig. 4). The basic data includes the name, address, telephone number, fax number, name and ID number of the responsible person, and uniform number of the subcontractor. The scoring of each evaluation item is conducted based on the basic data.

4.2 Quotation Module

The .Quotation Module. provides the data of the subcontractor and the project, the settings of each quotation item and the quotation data. (Fig. 5). This system also provides an enquiry function for the user to enquire relevant data of the subcontractor.

When selecting the subcontractor and the project, the user may select the work to be quoted and input the selection in the field of Quotation Data (Fig. 6).

4.3 Subcontractor Selection Module

Price is usually the key criterion for the traditional subcontractors selection. However, this traditional price-oriented selection approach only emphasizes the price and neglects the quality, timeframe and other factors supposed to be considered during the procurement process [6]. In the period of recession, this approach may lead to an awarding price that is much lower than the average bidding price, bringing a higher risk of failure to the owner [7]. Since each construction company has different consideration for each project. People may want to adjust the weights for each evaluation factor to reflect the real need. This study, therefore, conducts the analyses simultaneously for each evaluation item. Different weights can be set depending on the demand of each project.

The .Subcontractor Selection Module. provides the functions of project selection, setting of evaluation criteria, selection of subcontractors and awarding operation. The user is requested to choose a project and select the item to be evaluated from the options of the project. The system will arrange the subcontractors that have provided the quotation for the selected item. The user may then compare the quotation based on the data provided by the system (Fig. 7). During the awarding operation, the system will arrange the data to be compared and provide it to the user for the selection of subcontractors. When the awarding operation is completed, the system will calculate the awarding progress and aggregate the awarding amount.

4.4 Budget Control Module

The .Budget Control Module. provides the functions of project data selection and project quotation enquiry. When the awarding budget exceeds the planned budget, the system will give an alert to users (Fig. 8).

4.5 Report Preparation

All reports related to the procurement can be printed by using the .Report Preparation. function. The Progress Chart of Procurement can be printed by using the sub-function of the .Report Preparation. (Fig. 9).

4.6 Maintenance of Relevant Data

This function provides the maintenance of unit data, owner's data, project data, basic database, resources data and selection data. The evaluation criteria and their weights are created by the system based on the calculation of Section 3.3. To provide scalability to meet the future demand, the system allows the user to input the data to be changed here.

V. Conclusion

The results of the study are described in the following two major points:

5.1 Creation of Evaluation criteria and Their Weights

The study analyzes the subcontractor selection operation and uses AHP to calculate the weight of each selected evaluation item. The evaluated items and their weights are then integrated into the aided system to select appropriate subcontractors under the consideration of the price. The project may be adjusted to meet actual demand. The evaluation criteria and their weights acquired in the study are shown in Table 1.

5.2 Development of the Subcontractor Selection System

The study discovers the shortcomings of the current procurement operations and proposes feasible management strategies. The developed aided system provides the functions of the maintenance of subcontractor's data, quotation module, subcontractor selection module, budget control module, report preparation and maintenance of relevant data, which are capable to provide the following services:

A. The system sets the tasks for selected project and create unit price analysis and resources management data for effective cost analysis and control.

B. The system can select appropriate subcontractors based on the result of the evaluation. A fair and open process is provided by the system for construction companies to select appropriate subcontractors objectively.

C. The system successfully combines the procurement operations with the budget control, and is capable to give an alert when the budget exceeds its range.

D. The system provides the progress control function for the procurement operations, so that the decision makers may investigate the progress.

E. The objectivity of the evaluation criteria and their weights is confirmed by the industry. Both the evaluation item and the weights can be adjusted depending on the demand of each project.

VI. Acknowledgement

I would like to take this opportunity to show my appreciation to National Science Council, Executive Yuan, for its support of the study. Project No.: NSC 90-2211-E-216- 015.

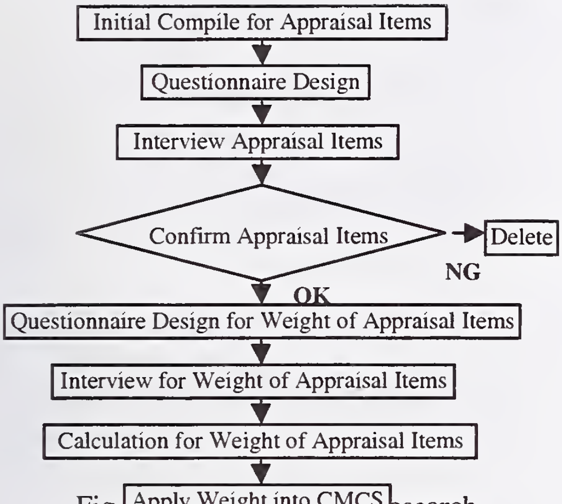


Fig. esearch

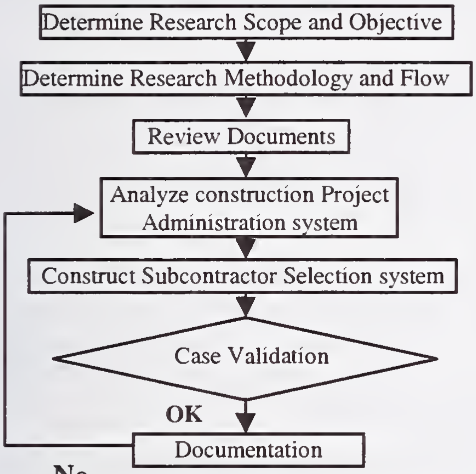


Fig. 2 Flowchart of Evaluation Criteria

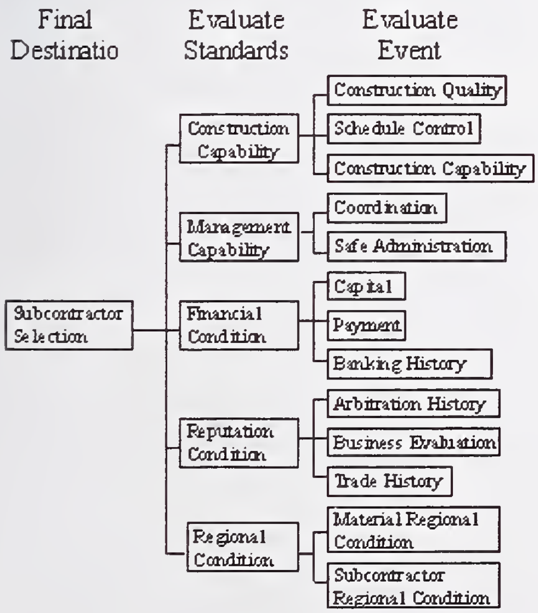


Fig. 3 Hierarchical Structure of the Factors for the Selection of Subcontractors

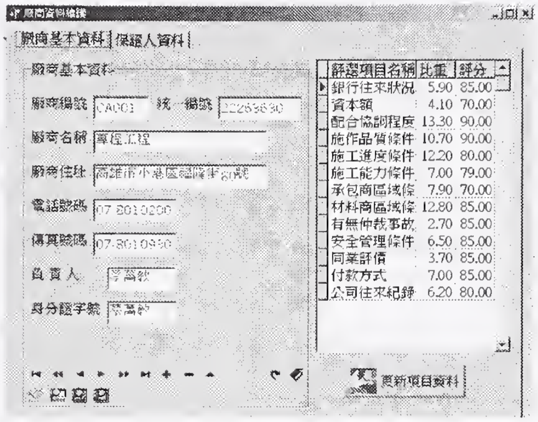


Fig. 4 Data Maintenance. Dialog

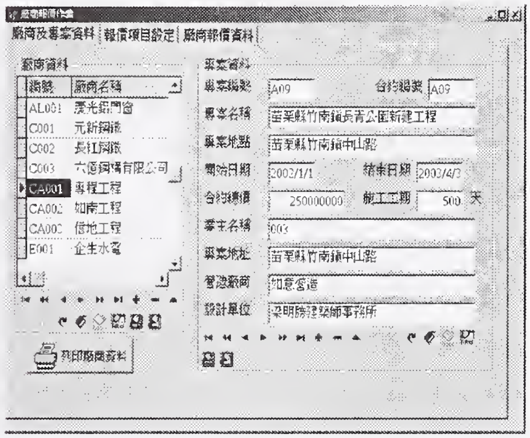


Fig. 5 .Quotation Module. Dialog

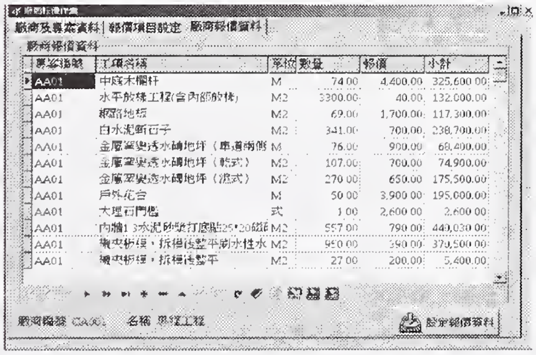


Fig. 6 .Quotation Data. Dialog

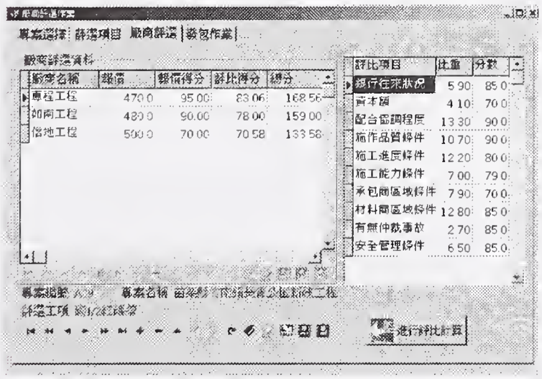


Fig. 7 Subcontractor Selection Screen

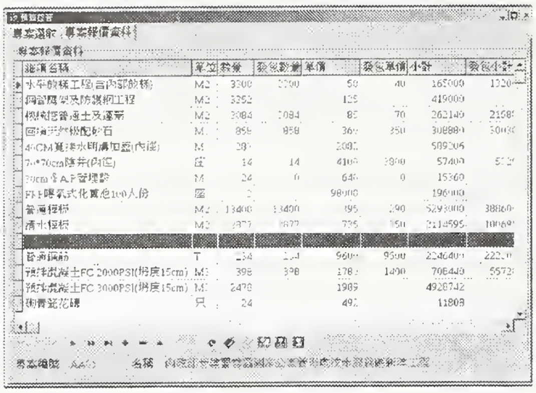


Fig. 8 Budget Control Screen

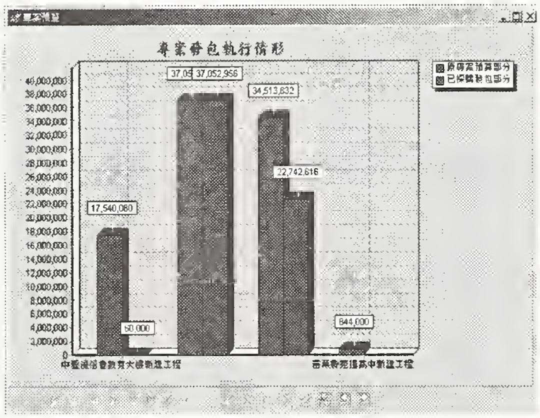


Fig. 9 Progress Chart of Procurement

Table 1. Relative Weight of the Evaluation Criteria

Evaluate Standards		Weight
Construction Capability	Construction Quality	0.107
	Schedule Control	0.122
	Construction Capability	0.070
Management Capability	Coordination	0.133
	Safe Administration	0.065
Financial Condition	Capital	0.041
	Payment	0.070
	Banking History	0.059
Reputation Condition	Arbitration History	0.027
	Business Evaluation	0.037
	Trade History	0.062
Regional Condition	Material Regional Condition	0.128
	Subcontractor Regional Condition	0.079

VII. REFERENCES

1. Ji-Zhong Zhang, "Research on Construction Subcontract Management", Master Dissertation, Department of Science and Construction Engineering, National Taiwan Institute of Technology, Taipei, pp141~ 151, 1996.

2. Zhong-Fa Huang, "Research on Construction Subcontract Management of Tactic and Organization", Doctoral Dissertation, Department of Civil Engineering, National Central University, Tao-Yuan Taiwan, pp126~131, 2000.

3. Satty, T.L., "The Analytic Hierarchy Process", McGrawHill, New York, 1980.

4. Zhen-Yuan Deng And Guo-Xiong Zeng, "Application of Analytic Hierarchy Process (AHP) on Intension Characteristic And Analytic I", National Statistics Report, pp5~22, 1989.

5. Zhen-Yuan Deng And Guo-Xiong Zeng, "Application of Analytic Hierarchy Process (AHP) on Intension Characteristic And Analytic II", National Statistics Report, pp1~15, 1989.

6. Zohar Hersman and Ralph Ellis, .Multi-parameter Bidding System-Innovation in Contract Administration. Journal of Construction Engineering and Management, ASCE, Vol.118, No1, pp142-150, 1992.

7. Gysbert J. R. Van Der Meulen and Arthur H.Money..The Bidding Game,.Journal of Construction Engineering and Management, ASCE, Vol.110, No.2, pp153- 164, 1984.

Development of Electronic Acquisition Model for Project Scheduling (e-AMPS) Using Java-XML

Will Y. Lin

Ph. D. Candidate, Dept. of Civil Engrg., National Taiwan University, Taipei 10617, Taiwan, E-mail: will@ce.ntu.edu.tw

Abstract: During the construction phase, participants in a multi-contract project acquire external real-time scheduling information from other involved parties and use this to make appropriate decisions in regard to project control. There are two major obstacles to project participants gaining efficient access to external information in a distributed data environment: (1) the variety of data structures that project members may use, and (2) lack of an automatic mechanism for data acquisition. Based on the ontology defined by eXtensible markup language Schema (XML Schema) and an automatic mechanism called Message Transfer Chain (MTC), an Electronic Acquisition Model for Project Scheduling (e-AMPS) centralized in an information agent, Message Agent (MA), was developed. Each participant equips a Message Agent as his unique information window to automatically acquire external information and provide other participants with scheduling information as well. The ultimate goal of this study is to build an automatic communication environment for multi-contract projects to solve the abovementioned difficulties, and thus achieve effective communication among project participants.

Keywords: Internet; WWW; Project Management; Intelligent Agent; Information Integration

Introduction

Most scheduling theories take into account a variety of situations, such as weather, site layouts and so on, according to available information while scheduling to produce a "perfect" schedule, which seems to forecast the future very well. However, due to a large number of construction uncertainties before the project starts, such as material shortages and interference between two tasks of different subcontractors, it is common that the initial schedule has such a variance with the real condition of construction that some planned tasks cannot be carried out accordingly. Recent planning-related research, such as Lean construction suggests that schedules should be updated adequately and constantly after the construction starts, according to the real-time engineering information available to keep themselves concurrent and useful. In most multi-contract projects, however, it's common for 80-90% of the tasks to be performed by subcontractors such that scheduling for these projects is a cooperative task which requires many project members to take part in. In order to realize the continuous scheduling suggested by Lean construction under this circumstance, it's necessary for these subcontractors to "dynamically communicate" together.

Communication in construction industry during construction phase is extremely complex. In terms of information technology, communication can be simplified as the exchange and reuse of information or messages between two independent parties. In this sense, to automate the communication among construction project members implies to automate the exchange and reuse of information or messages. The exchange and reuse of engineering information have been an issue in the field of automation in construction since information technologies were first introduced in 70's. Much research and related applications have also been developed to achieving all kinds of automation in communication. However, there still are two major obstacles to automate the continuous and collaborative scheduling for multi-contract projects: (1) the variety of data structures for scheduling that project members may use, and (2) the lack of an automatic mechanism for data acquisition in such a multi-user workplace for most multi-contract projects.

Based on the ontology defined by the eXtensible markup language Schema for Scheduling (XSS), the Data Acquisition Language for Scheduling (DALS), the Hierarchy Searching Algorithm (HSA), and an automatic mechanism called Message Transfer Chain (MTC), an Electronic Acquisition Model for Project Scheduling (e-AMPS) centralized in an information agent, Message Agent (MA), was developed. Each participant equips a Message Agent as his unique information window to automatically acquire external information and provide other participants with scheduling information as well. The ultimate goal of this study is to build an agent-based communication environment for multi-contract projects to solve the above-mentioned difficulties in automating communication in a multi-user workplace, and thus realize continuous and collaborative scheduling.

Architecture of e-AMPS

To solve the difficulties involved in sharing scheduling information among project participants in a data-distributed environment, an agent-based communication environment called Electronic Acquisition Model for Project Scheduling (*e*-AMPS) has been developed. The model is centralized in an information agent called Message Agent. Basically, Message Agent is a computer program that deals with all messaging tasks involved in automatic communication, and will be introduced in the following sections. Each participant in the same project, named a *Host* or *Contact Node* in the following paragraphs, equips a Message Agent as a unique information window so that Message Agents in the same project can automatically communicate with each other. In this section, we introduce the basic framework of *e*-AMPS and the functions of Message Agent in order to give an overall picture of the proposed concepts. Figure 1 illustrates the complete architecture of e-AMPS. The complete automatic communication consists of two different levels of replying to the imported requests: Data-retrieving level and Decision-making level. In this paper, we only focus on the Data-retrieving level. However, the components within the Decision-making level are also addressed to some extent in this section to help draw a more complete picture of our model. The complete framework of e-AMPS consists of five major components: Ontology base, Message Agent (MA), Open Data Repository, DALS-speaking Decision Support Systems for Scheduling, and Message Queues [1].

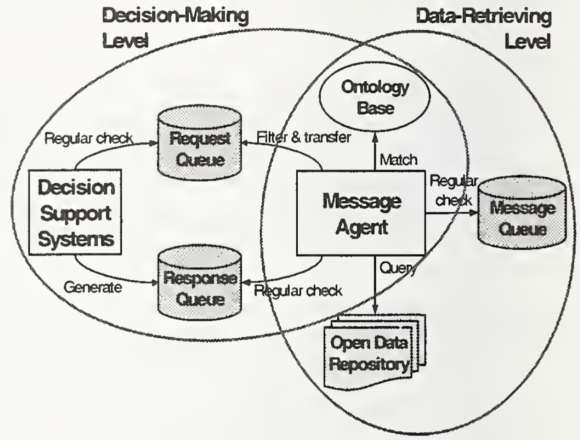


Figure 1: Main Framework of e-AMPS

(1) Ontology base: This stores all ontology on scheduling in terms of data schema, called XSS, the syntax of which is adopted from XML schema in our study. The ontology here is defined as "a specification of a conceptualization, or a description of the concepts and relationships that can exist for an agent or a community of agents."

(2) Open Data Repository: This contains scheduling information with standard data structure defined by the ontology (XSS), whose data structure is shown in Figure 2, and is in XML syntax [1]. It's basically a file folder that contains all scheduling information files of standard formats. There are two kinds of scheduling information files for each e-AMPS: Schedule File (schedule.xml) and Contract File (contract.xml).

(3) Message Agent: This deals with all manipulation of incoming and outgoing messages following the communication mechanism built by e-AMPS concepts. It communicates with other Message Agent mounted on other contact nodes, and also with its local decision support systems through the mapping table.

(4) DALS-speaking Decision Support Systems for Scheduling: They are built by the host, independently from the Message Agent. They have independent decision support models to

generate specific decisions toward certain fields. Most important of all, these decision support systems all recognize the DALS and use it to request for information from other project participants as their input data [1].

(5) Message Queues: Message Queue, physically an open access file folder, contains all messages (requests or responses) from other Message Agents. Each Message Agent will access its Message Queue regularly and automatically to react according to the messages [1].

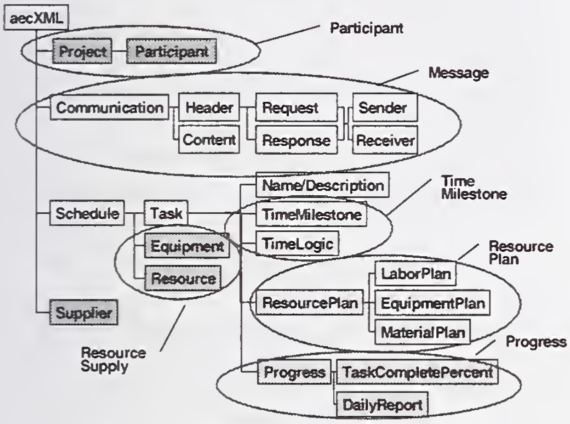


Figure 2: Data Structure of XSS

The utilization of e-AMPS can be divided into three stages:
(1) Installation: Every practitioner in the construction industry, i.e. Owner, A/E, contractor, and supplier, equips a Message Agent that is implemented by Java, regardless of platform to be used, and therefore a communication environment is then built where Message Agents in the same project can automatically communicate with each other. When installing the system, a practitioner is prompted to specify the local file folders for the Ontology Base that contains the data schema file, the Open Data Repository that contains the scheduling files, and the Message Queue that contains the message file. The folder for Message Queue should also be associated with a Uniform Resource Indicator (URI), which can be openly accessed by Message Agents of other project participants.

(2) Contract signing: Once the practitioner signs a contract with another party, the scheduling information is then prepared according to the data schema, and is deposited in the Open Data Repository, where the Message Agent can access and make inquires about the shared scheduling information.
(3) Contract execution: While the contract is being carried out, the Message Agent checks the Message Queue regularly and deals with all messaging tasks automatically according to the message it receives, such as sending the requests originating from the Host or passing the responses sent by other Message Agents to another Message Agent.

The detailed design of the Message Agent can be referred in another paper [1].

Implementation of Message Agent 1.0

Java ™ language is a rich environment for XML programming since there have been more XML-specific resources available in Java than in any other programming language. There are two major reasons why Java meets XML programming. The first is their shared reliance on the Internet. XML was designed to be straightforwardly usable on the Internet, while Java was designed to be used over the Internet. Java works well in a distributed environment, allowing users and programs to share information easily, while XML provides a tool for distributing and storing that information. The other reason is their shared use of hierarchical structures. Java's object-orientation and XML's fundamental use of nested hierarchies is a suitable match of combination. Programmers can easily develop tree structures with Java that match the structures of an XML document, making it easy to convert XML files into instantly usable data in Java application or applet. Due to the above reasons, this research uses Java 2 as the developing language. The Java ™ API for XML processing has been added to the Java 2 Platform. It provides basic support for processing XML documents through a standardized set of Java Platform APIs, and other network-specific programming facilities suitable for the implementation of e-AMPS and Message Agent. Several programming features are addressed first, which are multi-thread processing, parsing with a validating mode using XML Schema, and the use of Remote Method Invocation (RMI).

(1) Multi-thread processing: A thread — something called an execution context or a lightweight process — is a single sequential

flow of control within a program. A single thread process means that a process has a beginning, a sequence, and an end and at any given time during the runtime of the thread, there is a single point of execution. On the other hand, multiple threads mean that there are more than one single thread running at the same time and performing different tasks within a program. Since carrying out various manipulations of a message, the Message Agent is implemented with multiple threads and thus different manipulations of a message are able to proceed independently and smoothly.

(2) Validating documents using XML Schema: There are two types of XML parsers, divided by different function levels: validating parser and non-validating parser. There are also two methods to validate an XML document: using Document Type Definition (DTD) or using Schema. An XML document is valid if it has an associated DTD or Schema, and if the document compiles with the constraints expressed in it. A DTD defines the data structure of an XML document. It specifies the order in which tags occur, what the tags are, and how many tags are allowed. A DTD provides a uniform format for defining the structure and markup of an XML document. Unlike DTDs, however, XML Schemas adhere to the XML specification and provide better support for XML namespaces and more data types. It is also a recommendation of the W3C. Schemas provide a more flexible means for defining the structure, content, and semantics of XML than DTDs. In many areas of application, DTD is replaced with XML Schema nowadays although DTDs had been widely adopted for years.
Due to the above-mentioned advantages of XML Schemas, the Message Agent adopts a validating parser using XML Schema.

(3) Use of Remote Method Invocation (RMI): Since several major manipulations of a message are involved in passing an XML-based message from a local Message Agent to remote Message Agents, an approach of file transferring from one host to another is required by the Message Agent. Although the protocol File Transfer Protocol (FTP) or other message transfer methods such as SOAP is a possible way to be applied to this end, the Message Agent 1.0 adopts a remote access mechanism provided by Java called Java Remote Method Invocation (RMI) since RMI allows an object running in one Java Virtual Machine (VM) easily to invoke methods on an object running in another Java VM. RMI provides for remote communication between programs written in the Java programming language.

Figure 3 is the flowchart of starting up and stopping Message Agent, which is the main stream of the whole program. Since the process of dealing with messages undertaken by the Message Agent is a routine task with a given running period, the main stream starts at arousing a thread called ***MainLoop()***. ***MainLoop()*** then triggers three child threads: ***MessagingTask_AppendingMessage()***, ***MessagingTask_CheckingMessage()***, and ***MessagingTask_DispatchingOutboxMessage( )***.

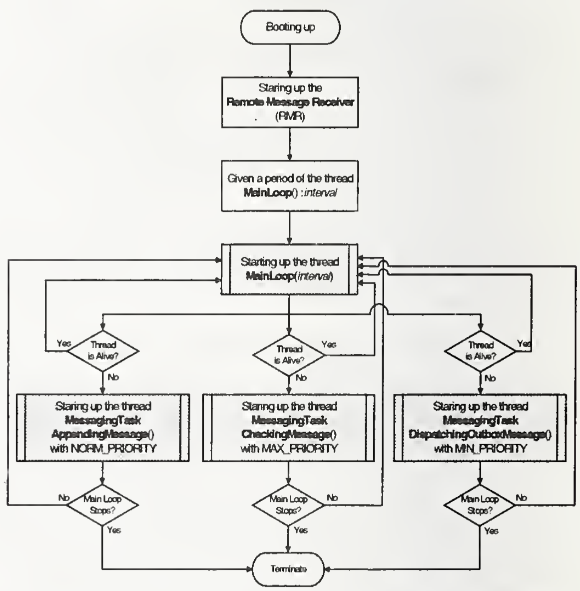


Figure 3: Flowchart of Running Message Agent

In three child threads of the thread ***MainLoop()***, whether the thread lifecycle of last execution is finished or not would be examined first. If the thread is still "alive", the new thread will not be triggered. Thread priority is set mainly according to the average running time spent. The thread that spends the longest time averagely gets highest priority. Figure 4 illustrates the core objects of Message Agent version 1.0 and their relationships one another. The class ***MainFrame*** is the visual user interface that initiates the root class ***MessageAgent*** of the whole program. Under the root class, there are thread class ***MainLoop*** with three child threads, five major message processor/manipulation classes, and a **RMI** class ***RemoteReceiver*** that implements the interface ***FileReceiver***.

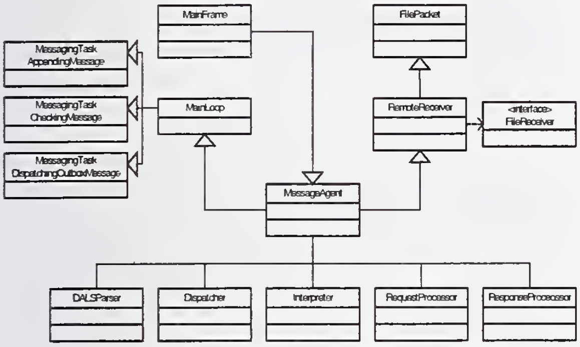


Figure 4: The Relationship between Objects

Scenario

A hypothetical design-build project is made up to illustrate more fully the concept of e-AMPS and the effect of the Message Agent. The milestone network and bar chart of the project are shown in Figure 5.

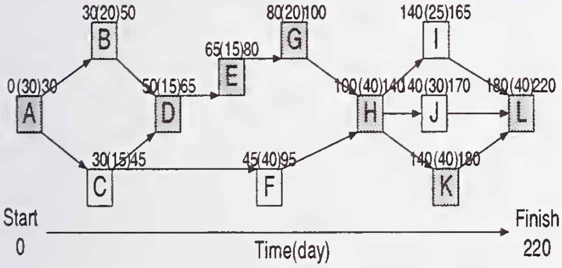


Figure 5:The Project Network of the Example

There are 12 packages enclosed in this project, undertaken by a general contractor and his 11 sub-participants, from P1 to P11. Figure 6 shows summary bar charts of all sub-participants under the cooperation structure of the example project. The entire project starts on Jan 1st, 2002 and finishes on Aug 8th in the same year using a calendar of 7-workingday a week due to simplify the complexity of the example.

In following paragraphs, a scenario is made up respectively associated with a typical communication behaviors for scheduling: requesting for progress data. The whole communication cycle, from the original request to the terminal responses, is recorded and represented as well as some important facts and results are extracted to emphasize the effect of e-AMPS.

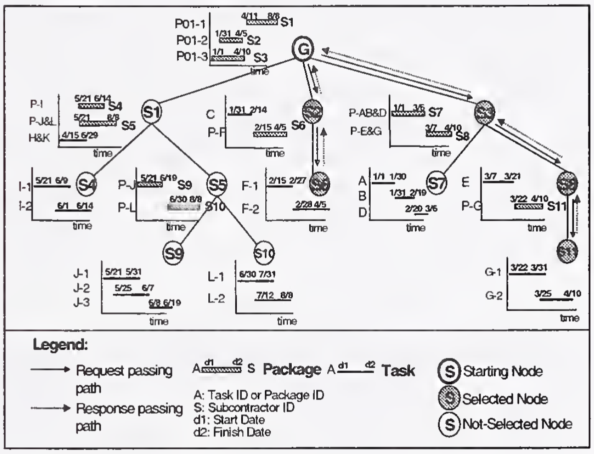


Figure 6: Message Passing Path of the Scenario

On Mar 10th 2002, the general contractor originates a request for the progress information of all tasks whose duration overlaps a period between Mar 15th and Mar 25th, about two weeks before the **Task H&K** of participant P1 starts. The original request without the header created by a pre-programmed process at the general contractor's site is deposited in the Message Queue, and waits for his Message Agent to dispatch it. The Message Agent of the general contractor detects this request and automatically performs the HSA. Since having no upper messengers, Message Agent of the general contractor decides to dispatch the request to two of its lower messengers, S2 and S3, since the packages undertaken by S2 and S3 meet time and scope constraints.

The request to S2 is bypassed to S2's lower, S6, due to Package P-F undertaken by S6 meets the constraints specified by the request. Upon receiving the bypassed request from S2, S6's Message Agent perform the query transformation and generate a response sending back to S2, flowed by another bypassing by S2 back to the general contractor.

Figures from Figure 7 to Figure 9 show the sequences of message manipulations by the Message Agents of participants involved in this scenario, G, S2, and S6, respectively. At the site of G in this scenario, G's Message Agent passes two requests (requestId: **3945** & requested: **8379**) to two of his lowers, S2

(URI: 140.112.10.31) and S3 (URI: 140.112.10.32), respectively. 35 seconds later, it receives the first one response (responseId: **3884**) from S3 in which the original replier is S8, according to the message log at the site of G. 10 more seconds later, it receives the second responses (responseId: **3389**) from S2, in which the response is generated by S6. 6 more seconds later, it receives the last response again from S3, in which the original response is generated by S11 and is bypassed through S8 and S3 in turns.

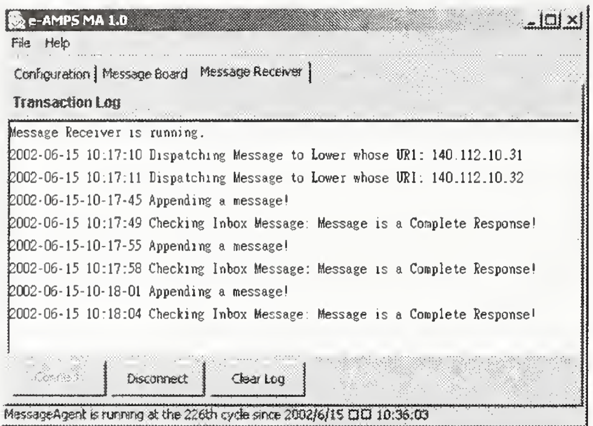


Figure 7: Message Manipulation at the site of G (URI: 140.112.10.16)

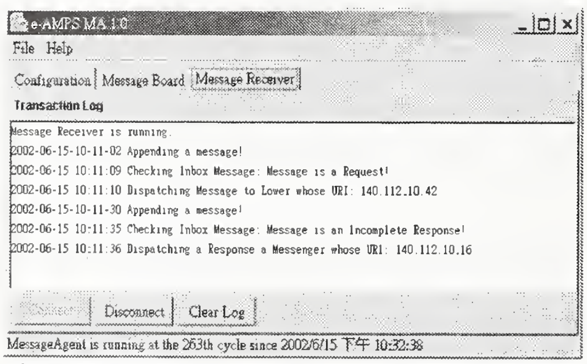


Figure 8: Message Manipulation at the site of S2 (URI: 140.112.10.31)

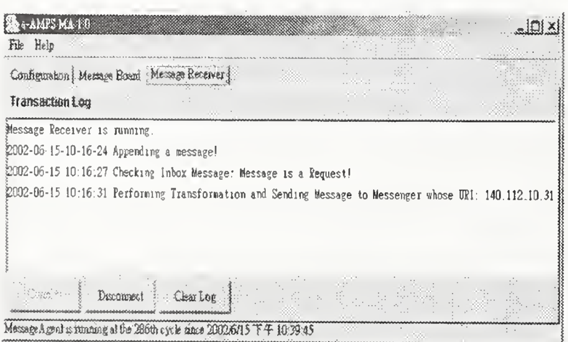


Figure 9: Message Manipulation at the site of S6 (URI: 140.112.10.42)

Conclusion

Sharing of project scheduling information among subcontractors is useful for predicting potential delays and taking any necessary precautions. However, there are two major obstacles to multi-contract project participants accessing the external information they need efficiently: (1) the variety of data structures that project members may use, and (2) lack of an automatic mechanism for automatic data acquisition. An agent-based communication environment called Electronic Acquisition Model for Project Scheduling (e-AMPS) is developed to solve the abovementioned shortcomings. Message Agent was implemented using Java 2 and tested in IBM PC with Windows 2000 OS. The testing and system performance have been evaluated with positive results.

Acknowledgement

The authors would like to acknowledge the National Science Council, Taiwan, for financially supporting this work under contract No. NSC-90-2211-E-002-074.

Reference

1. W. Y. Lin, "Development of Electronic Acquisition Model for Project Scheduling (e-AMPS) Using Java-XML" PHD Dissertation, Department of Civil Engineering, National Taiwan University, 2002.

An Automated Selecting Subcontractor Model In E-Commerce

by

Pao-Hung Lin

Asst. Prof., Dept. of Civil Engineering, Feng-Chia University, Taichung, Taiwan

E-mail: paolin@fcu.edu.tw

Abstract: From the observation of organization changes in digital economy, the features of organization expansion from its concentrated kernel to the globalized e-commerce scope is identified to rebuild the relationships and connections between the whole subcontracting supply chain for construction projects. An automated subcontracting selection model is proposed and the final decision-making for selection will be approached by Multiple Attribute Utility Theory. Following this methodology, it can be expected to bring the excellent decision support into full play for the selection of subcontractors.

KEYWORDS: utility theory, e-commerce, subcontract, decision-making

1. INTRODUCTION

Knowledge-based and networked characteristics shape the all new information era in the digital economy. "Electronic Commerce" (e-commerce) has been the symbol of enterprise activities in digital time. The force of virtualization and virtual market constitutes the heart of e-commerce in the digital economy. The nature of e-commerce is to build the innovation of competitive advantage between enterprises by way of web connection, to alter the operation process and to enhance the efficiency of organization. Digitization not only improves the information quality and enables interactivity, but also provides a powerful approach for management in many respects. The new economy is creating a reformation of conflicting drivers causing every enterprise in every industry to rethink its competitive advantage, to rebuild its system of value chain, and to restructure its organization pattern. All the traditional thinking way to communication or transaction is facing the unprecedented challenges to shift.

Tapscott (1996) pointed that the result of this shift from physical to virtual does not simply reduce redundancy and speed up communication. Rather, the substantial changes in the essence of human and organizational communication occur. Inter-enterprise computing is causing new economic dynamics in the following ways: accessibility of partners, new interdependence, interorganizational metabolism, cooperative competitiveness, and interorganization value creation. Through the connection of information network, the resources are distributed and values are generated in the boundless virtual marketplace.

Due to the sharp differences between the "virtual market" transaction atmosphere in the post-information era and the traditional physically economic market, major impact are certain to occur within traditional main contractor and subcontractor relations. The processes of these change, and whether they will completely alter previous subcontracting relationships and models is a topic worthy of further research and consideration. From comparisons of the construction industry with the manufacturing industry or other productions that have gone through relatively little in the way of developing IT environment, it is quite clear that, based on the ideas of e-commerce and the importation of practical methods, in the future it will be possible to improve on abuses that arise in traditional construction industry business processes.

Although it's crucial to select appropriate subcontractors to implement specific construction project, few research really touched the selection techniques and methodologies and extended further into the connection of overall supply chain. Therefore, what the influence to the whole subcontracting supply chain of the research's concern in this digital economy is analyzed from the new vision of economics of electronic commerce. The new transaction mechanism is shaping in digital economy and all the trends and properties existing in the omnipresent environment have resulted in great change of organizational structure and the nature of transaction, and even paradigm shift.

2. METHODOLOGY

The methodology of this research is founded on the observation of e-commerce economics as well as networking mechanism and put the problem in center of subcontracting supply chain management or the combination of subcontractors. The major significance of this methodology is to support the acquisition of optimal selection set from myriad subcontracting combinations for specific construction project. All the bidding information and decision are implemented through omnipresent Internet. Every client user needs only to open their web browser to use the system's capabilities to retrieve all of the needed information online. The architecture of data transformation can be shown as Figure 1. Further, this methodology is proposed by taking advantage of the portfolio theory of investment choice under conditions of uncertainty on the mean and the variance of the distributions of returns to approach the selection model of subcontractors in construction project. That is, the risky connection in whole supply chain will be considered and obviously the profit will no longer be the only factor to the decision-making of bidder selection as the traditional bidding strategy. And most importantly, the factor that must be taken into consideration in electronic bidding system is how to provide a compatible environment and services meeting the needs of all participating enterprises of the global market in a timely and sharing manner.

This methodology also extends the decision manner of two factors of portfolio theory to multiple attribute decision-making in order to be endowed with potential for encompassing all the possible factors of influence to selection. The final selection will be dependent on the risk preference of decision-maker by the utility function. Comprehensively, the methodology is enclosed by the internet-based information technology as the demonstration of Figure 2.

3. NETWORKING MECHANISM

In retrospection of the development of commercial organization, there can be classified as function oriented, department oriented, and the matrix type in traditional organization categories (Miles and Snow, 1992). Nevertheless, "networking" has been the main research focus in management field and the trend to the application of reality. There are many different denominations to interpret the remold organization, such as Peter Drucker's (1998) "Networked Organization," Peter Senge's "Learning Organization," Keen's (1991) "Relational Organization" and Davidow and Malone's (1992) "Virtual Organization." They all put emphasis on the "relational" network between members of organization. No matter what the idea of networking interpreted from the view point of resource dependence (Joanson and Mattsson, 1987), transaction cost (Williamson, 1975; Oliver, 1990), or strategic alliance (Porter and Fuller, 1987), the essence of networking lies on that the specific enterprise can concentrate on the most valuable activities with competitive advantages, to reach the economies of scale and obtain the benefit of work specialization (Porter, 1985). The "networked organization" opens a new style and feature in the area of organization structure and strategic management, revives the traditional organization pattern (Miles and Snow, 1992), and provides the alternatives except vertical integration as well as diversification.

From the investigation and observation of the evolutionary process in traditional organization relationship, it can be found that the "information network" is indeed playing a very crucial role to push the changes and virtualization (Tang, 1998). Figure 3 represents the process of organization development under both the coordinate axes of relation network and information network. It demonstrates the features of organization expansion from its limited kernel, partnership, vertical integration, outsourcing to the e-commerce global scope according with the developing information network, and simultaneously expresses the virtualized evolution of organization relationship. Also, the corresponding space-time situation of this research's concerns are located on the relation plane to show the focus of the present model. Additionally and importantly, this research is not only focusing on the exploration of the forming of virtual organization, but on the searching of the profit-base and efficiency to improve the link of selecting subcontractors by observing the organization changes in digital economy.

4. THE PRESENT CONCEPTUAL MODEL

The proposed selection model takes three main attributes into subcontracting consideration, namely, the expected return (*ERN*), the planned performance dispersion (*PPD*), and the comprehensive risk index (*CRI*). The expected return and the planned performance dispersion are the leading roles and have been exactly discussed by Tserng and Lin. This research, the comprehensive risk index is to integrate the bidder's credit degree and quotation deviation (including tender price and duration) to expansively consider the correlation between subcontractors in the specific combination of construction project. In the concrete, the *CRI* is a representative evaluation index composed by three kinds of uncertainty α_{ij}, β_{ij}, and γ_i, where α_{ij} is the measurement of dispersed degree of quoted duration for subcontractor i to the market value in the bidding of specific subproject j ; β_{ij} is the measurement of dispersed degree of quoted price for subcontractor i to the market value in specific subproject j ; and γ_i is the credit rating for subcontractor i evaluated by the general contractor in advance. Expressed in equation form and taken the absolute value in numerator to eliminate the negative situation, it shows that

$$\alpha_{ij} = \frac{\left| Q_{D_{ij}} - \mu_{D_j} \right|}{\sigma_{D_j}} \qquad (eq-1)$$

in which $Q_{D_{ij}}$ is the quoted duration of subcontractor i in the bidding of specific subproject j ; μ_{D_j} is the mean value of all quoted duration in bidding market for specific subproject j ; σ_{D_j} is the standard deviation of all quoted duration in bidding market for specific subproject j. In the same way, β_{ij} can be obtained. The present model denominates the comprehensive uncertainty of ($\alpha \bullet \beta \bullet \gamma$) as "quasi-$\sigma$", compared with the physical variance σ^2 of specific asset returns in portfolio theory. Thus, it can be expressed as following equation

$$\sigma_q = \alpha \bullet \beta \bullet \gamma \qquad (eq\text{-}2)$$

From portfolio theory, the comprehensive risk index *CRI* can be expressed as

$$CRI^2 = \sum_{i=1}^{n} x_i^2 (\sigma_q)_i^2 + 2\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j>i}}^{n} x_i x_j (\sigma_q)_{ij} \qquad (eq-3)$$

$$= \sum_{i=1}^{n} x_i^2 (\sigma_q)_i^2 + 2\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j>i}}^{n} x_i x_j \rho_{ij} (\sigma_q)_i (\sigma_q)_j \qquad (eq-4)$$

where the proportion coefficient x_i in portfolio theory, here represents the weight factor which is equal to the ratio of the quoted price of specific subcontractor i *or* j to the quoted summation of the whole selected combination; and ρ_{ij} is the coefficient of correlation between subcontractor i and j.

The present conceptual model combined three attributes is depicted as Figure 4 in the space constructed by *ERN*, *PPD*, and *CRI* three dimensions. Straightly, the concept model is adherent to the two-attribute model of portfolio selection to demonstrate it in 3-D space. However, it is quite different and much more complicated to find the solution between the utility function surface and nondominated frontier surface in 3-D space than it is in the 2-D space. As a matter of fact, the mathematical construction of utility function surface and nondominated frontier surface is a very difficult task. Therefore the present model will take the multiattribute utility theory to search for the finally optimal selection for decision-making.

5. DECISION-MAKING MODEL

As stated in relevant literature, multiattribute utility theory evaluates utility functions intended to accurately express a decision-maker's outcome preferences in terms of multiple attributes. It really reduces the complex problem of assessing a multiattribute utility function into one of assessing a series of unidimensional utility functions. Accordingly, the three concerned attributes of *ERN, PPD*, and *CRI* in the present model are individually appraised by the preference of decision-maker. From Keeney's (1993) theorem, under the assumption of mutual preferential independence and mutual utility independence of the three attributes, the utility function can be expressed as either the multiplicative function or the additive function.

That is, the following equations:

$$U_{obj} = [\prod_{i=1}^{3}(1 + kk_i u_i(x_i)) - 1]/k \quad , \sum_{i=1}^{3} k_i \neq 1 \quad (eq-5)$$

or

$$U_{obj} = \sum_{i=1}^{3} k_i u_i(x_i), \quad \sum_{i=1}^{3} k_i = 1 \quad (eq-6)$$

where U_{obj} is the objective function of utility; k is the scaling constant; and $u(x)$ is the individual utility function for each attribute. The three attributes in decision-making will also be normalized to satisfy with the condition that the summation of all weighted scaling constants is equal to one. Exactly, the additive utility function of Equation 2 is held in the present model to formulate the objective function:

$$\textit{Maximize} \quad U_{obj} = \sum_{i=1}^{3} k_i u_i(x_i), \quad \sum_{i=1}^{3} k_i = 1 \quad (eq-7)$$

The maximum utility value implies the optimal selection for the combination of subcontractors. The solution for Equation 3 is necessary to decide the scaling constants and individual utility functions. The method of determining weights in additive utility models includes multivariate linear (MLR) or non-linear regression analysis, analytical hierarchy process (AHP), direct decomposed tradeoffs and so on (Schoemaker and Waid, 1982). In general, it can be separated into statistical versus subjective approaches. For simplicity and practical purposes, the present operation on weights will thoroughly leave to the arbitrary setting by the decision-maker. More, the sensitivity analysis will be tested to evaluate the changing influences of each weight. Consequently, the procedures for assessing individual utility functions unquestionably become the most significant key point in the present model.

Utility theory assumes that a decision-maker can choose among the alternatives available to the individual to meet the maximum satisfaction degree in individual preference. This can be certainly implied the individual is aware of those alternatives and is capable of evaluating them through "utility function." Moreover, relative to a series of objectives it is assumed all information belonging to the various levels of the objectives can be captured by an individual's utility function. In effect, a utility function is simply a mapping of the values in the range of an attribute into a cardinal worth scale and can be stated as a formal, mathematical representation of individual preference structure. Thus if an individual conforms to specific axioms, a utility function can be constructed. Utility functions have been used extensively in consumer demand theory and economics, and they have been applied to private and public decision-making problems (Goicoechea et al., 1982). For multiobjective problems utility functions can completely order the set of nondominated objectives. The nondominated solution which yields the highest utility is referred to as the *best-compromise solution*.

6. CONCLUSION

The present model system is constructed under the internet-based environment. All the participants in this system are necessarily provided with the basic requirements and functions for web-usage. In other words, the initial investment to build up a suitable networking environment is indispensable to this system. Under the current developing environment, the testing case had ever been completely dealt with in a few minutes to provide decision-making support following this methodology. Although it is very difficult to evaluate the precisely quantitative benefit and economic efficiency between the traditional way to subcontract and the present methodology, it can be greatly expected to reduce the overall amount of time and capital required for the complicated process of procurement. Above of all, it apparently enhances the potential ability to maximize the preferred degree of decision maker's utility in a scientific manner. In spite that the system performance is highly dependent on the computing complexity of raising problems and it will also be compressed by the bandwidth capacity as well as the flow of web transportation thereupon, nevertheless, the present model indeed brings excellent decision support into full play for the selection of subcontractors.

REFERENCES

1. Davidow, W. H. and Malone, M. S., *The Virtual Corporation*, Harper Collins Publishing, N.Y., 1992.

2. Drucker, P. E., "The Coming of the New Organization," *Harvard Business Review*, Vol. 66, No. 1, 1988, pp. 45-53.

3. Goicoechea, A,, Hansen, D. R. and Duckstein, L., *Multiobjective Decision Analysis with Engineering and Business Applications*, Wiley Publishing, New York, 1982.

4. Johansson, J. and Mattsson, L. G., "Inter-organizational Relations in Industrial System: A Network Approach Compared with the Transaction-Cost Approach," *International Studies of Management and Organization*, Vol. 17, No. 1, pp. 34-48, 1987.

5. Keeney, R. L. and Raiffa, H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge University Press, 1993.

6. Keen, P. G. W., *Shaping the Future: Business Design through Information Technology*, Harvard Business School Press, 1991.

7. Miles, R. E. and Snow, C. C., "Cause of Failure in Network Organizations," *California Management Review*, Vol. 35, 1992, pp. 53-72.

8. Oliver, E. C., "Determinant of Inter-organizational Relationships: Integration and Future Directions," *Academy of Management Review*, Vol. 15, No. 2, 1990, pp. 241-265.

9. Tang, H. C., *Study of Construction of Virtual Organization Model From Network Perspective, Example: the Electronics and Information Industries*, Ph.D. Dissertation, Graduate School of Business, National Taiwan University, 1998.

10. Tapscott, D., *The Digital Economy: Promise and Peril in the Age of Networked Intelligence*, McGraw-Hill, N.Y., 1996.

11. Tserng, H. P. and Lin, P. H., "An Accelerated Subcontracting And Procuring Model for Construction Projects,"

12. *International Journal of Automation in Construction*, 2002, pp. 105-125.

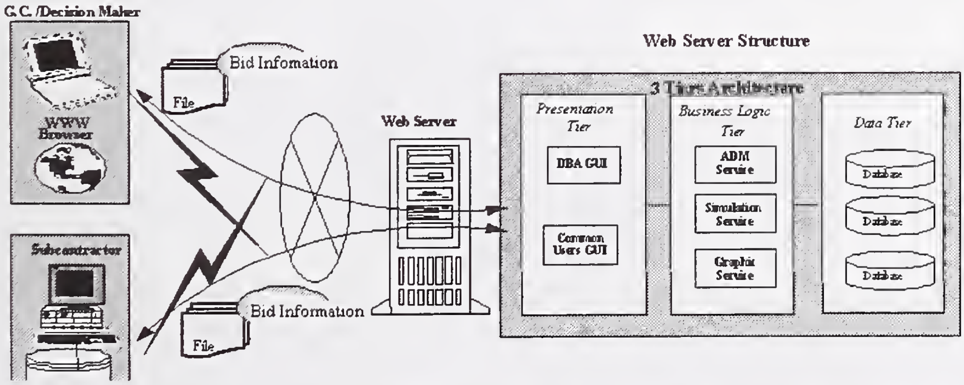


Figure 1. Architecture and Data Transformation of the Present Model

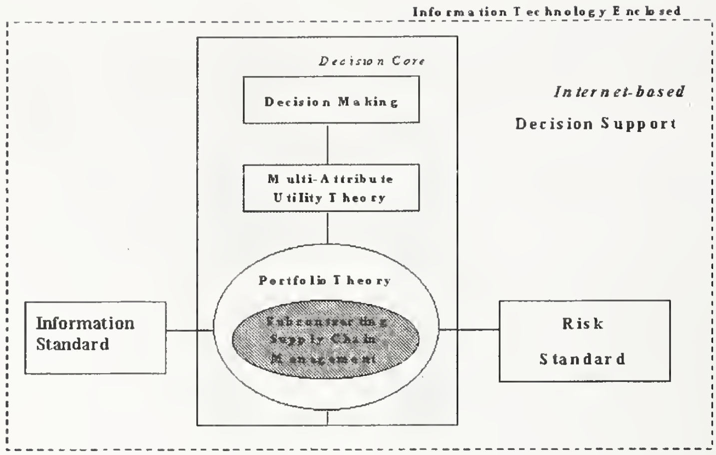
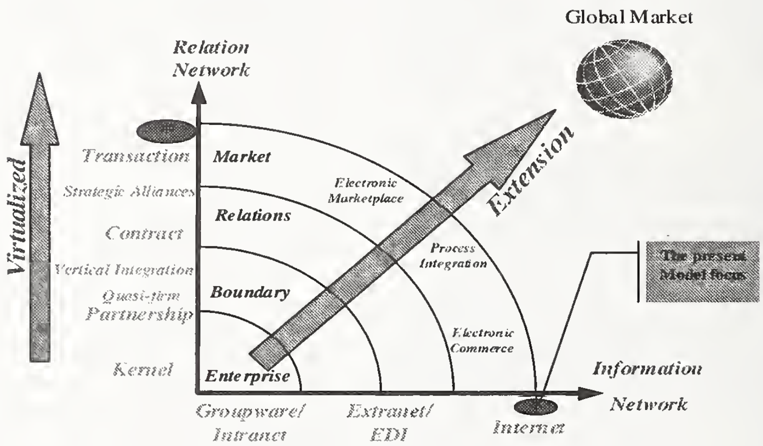


Figure 2. Methodology of The Present Model



Note: Modified from Konsynski (1993)

Figure 3. Evolution of Organization Type Under Networking Mechanism

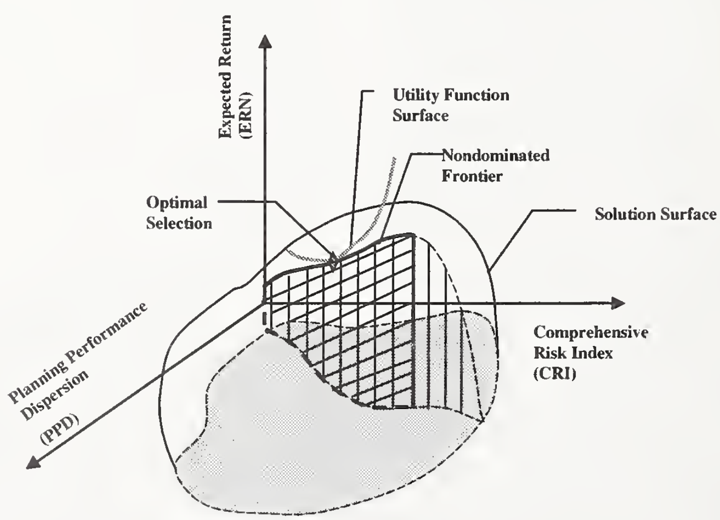


Figure 4. The Expression for Conceptual Model

MOBILE 3D VISUALIZATION FOR CONSTRUCTION

Robert R. Lipman[1]

ABSTRACT: Until recently, visualization of 3D models required computational resources available only on desktop computers in office environments. With the advance of technology, it is now possible to visualize substantial 3D models on mobile handheld computers in the field. This paper discusses some of the current technology, discusses the use of the Virtual Reality Modeling Language on mobile handheld computers, shows several examples of 3D structural steelwork models visualized on a mobile handheld computer, and identifies some of the limitations imposed by current technology.

KEYWORDS: construction; handheld computer; visualization; VRML

1. INTRODUCTION

Over the past decade, CAD software for construction has gone from simple 2D drafting programs to complex 3D solid modeling systems that support structural analysis, logical product models, project management information systems, collaborative environments, and task scheduling. The hardware on which those programs run has progressed from mainframes to desktop computer and laptops. The size of computer storage and memory has gone from megabytes to gigabytes. Communication between client-server applications has advanced from dialup modems to always on broadband and wireless networks.

Mobile handheld computers and personal digital assistants (PDA) such as PocketPC [1] and Palm [2] devices are now gaining acceptance as useful tools at a construction site [3]. Software applications previously confined to the engineer's office, are now available on the construction site. They provide a means to have access to project information that is stored locally on the PDA or accessed from a server through a wireless Internet connection. The information can also flow in the opposite direction where the PDA is used to collect data from the construction site and is communicated back to desktop applications for subsequent processing.

Although 3D solid modeling is becoming more popular, there are very few applications for doing 3D visualization on PDAs. A common format used for displaying 3D models on desktop computers is the Virtual Reality Modeling Language (VRML) [4,5], which is a scene description language for representing 3D interactive models on the web. VRML models are displayed in freely available VRML plugins [6,7,8,9] for Netscape or Internet Explorer. A VRML browser has also been developed for the PocketPC [10]. Most CAD programs can export their models in VRML format; however, the VRML representation of a CAD model is not well suited for viewing on a PocketPC because of its memory and computational limitations. The VRML generated from CAD programs usually does not take advantage of any of the useful features in VRML to make the file sizes smaller and more efficient.

2. HARDWARE

The most common type of handheld computers are PDAs that run either the Palm or Windows CE operating system. PDAs typically have a screen resolution of 240 x 320 pixels. This low

[1] National Institute of Standards and Technology; Building and Fire Research Laboratory; Gaithersburg, Maryland 20899-8630 (301) 975-3829 robert.lipman@nist.gov http://cic.nist.gov/

DISCLAIMER: Any mention of commercial products does not imply recommendation or endorsement by NIST.

screen resolution limits the type of 2D or 3D information that can be effectively displayed. Traditional methods of displaying data on a normal sized computer monitor are often not appropriate for PDAs and new visualization paradigms need to be developed. Devices running the Palm OS usually have 8 MB of memory while newer PocketPCs running Windows CE typically have 64 MB of memory. Depending on the operating system, the memory is used to store the operating system, applications, and files and provide memory to run programs. PDAs normally do not have a fixed hard disk. Currently the processors in PDAs are not very powerful or suited to dealing with large amounts of data that might be needed for 3D CAD applications. The first 2D graphics accelerators for PocketPCs are only beginning to be available [11].

The user interacts with the PDA using a stylus to pick menu items or to write on the screen. Some PDAs have a small joystick-like device to aid in navigation. Detachable keyboards can also be used to enter information. Files and applications are usually transferred to and from a PDA by synchronizing with a host computer. Information can be transferred directly between PDAs or other computers by beaming it through an infrared port.

Most PDAs have expansion slots for compact flash memory, wireless Internet access, and other CF or PC expansion cards such as global positioning system receivers or barcode scanners. Some PDAs have built-in wireless access. There are even PC expansion cards that incorporate a 1 GB miniature hard drive [12]. With this amount of storage on a handheld device, an enormous quantity of information can be brought to the construction site.

One of the primary limiting factors for displaying 3D data on a PDA is the amount of built-in memory. Adding memory or a hard drive through an expansion slot only increases the amount of space available for storing programs and files. The memory available to run programs is restricted to whatever is built-in to the PDA. For example, in a PocketPC running Windows CE with 64 MB of memory, by default, half of the memory is allocated to storing programs and files and the other half to running programs. The memory allocation between storing and running programs can be adjusted to provide more or less of the built-in memory to each function.

There are also several PDAs that run the Linux operating system. Larger pen-based tablet computers run the Windows XP operating system. One manufacturer has even developed a pocket-sized computer that runs Windows XP with a 10 GB hard disk and 256 MB of memory [13].

3. SOFTWARE

Several CAD and construction related software applications are commercially available for PDAs. In general, the more complex and graphics oriented applications run on PocketPCs with the Windows CE operating system. Applications that run on Palm devices are usually text oriented such as punch list, project management, or scheduling software. Research into collaborative dynamic project management has shown how PDAs and even web-enabled cellphones can be used as wireless display devices for traditional desktop construction management software [14].

A number of 2D graphics applications that display traditional CAD drawings in DWG, DXF, or DGN formats are available for PDAs. Some of the programs can display the drawings directly, while others require that the drawing be converted to a proprietary format that is optimized for the software. Other applications provide the means to markup existing or create new CAD drawings [15,16]. Certainly the low screen resolution and small screen size are limiting factors for the utility of working with CAD drawings on a PDA. Version control of CAD drawings between a PDA and desktop computer is also of concern. Simple file synchronization may be inadequate when format conversions are taking place on one platform but not the other.

The development of 3D visualization programs is only in its infancy. At least two commercial products are available for developing and delivering 3D content and applications for PDAs and other mobile devices [17,18]. However, these applications are geared towards 3D games and not to 3D visualization for CAD. Preliminary research at IBM led to the development of a 3D viewer that could display a subset of VRML [19] on a PocketPC.

Pocket Cortona [10] is an inexpensive VRML viewer for PocketPCs that supports most of the VRML specification including application specific nodes. VRML models that are generated by CAD programs can be displayed in Pocket Cortona without modifications. Pocket Cortona can run in a standalone mode or as a plugin to a web browser. In the standalone mode, the VRML file to be displayed is stored locally on the PocketPC. PocketPCs with a wireless Internet connection, web browser, and Pocket Cortona can access VRML models through a web page in the same way they are accessed on a desktop computer.

4. VRML

VRML is an ISO standard [20] for describing 3D geometric objects and their behaviors. Information in a VRML model is arranged in a scene graph that defines the parent-child relationship between VRML nodes. Some common VRML nodes describe 3D shapes, the appearance of the shapes, and geometric transforms that can be applied to the shapes. The shapes include primitives such as boxes, spheres, cones, cylinders, extrusions, and general sets of points and polygons known as IndexedFaceSets. Transforms that can be applied to the primitives include translation, rotation, and scale.

Typically the VRML models exported by CAD programs use IndexedFaceSets and geometric transforms to represent all of the geometry in a model. While this provides an accurate visual representation of the geometry, it does not take advantage of several features of VRML to reduce the file size and make the processing of the VRML model much more efficient.

The VRML prototype mechanism allows for the creation of application-specific geometric objects similar to the built-in primitives such as boxes, spheres, cones, and cylinders. Research at NIST [21,22,23] has developed a mapping between the CIMsteel Integration Standard (CIS2) [24] and VRML. CIS2 is a logical product model [25] for describing steel structures and has been adopted by the American Institute of Steel Construction as their standard for electronic data interchange [26]. VRML prototypes have been developed that correspond to CIS2 entities such as parts (beams, columns, braces, clip angles, plates), bolts, holes, welds, and locations.

Another important feature of VRML, which is usually not implemented in the VRML model exported by a CAD program, is the ability to define one geometric object and create many instances of it. In a typical steel structure there are many identical parts, each with their own location. In the VRML model only one of each type of part needs to be modeled. The rest of the VRML model is built from instances of the defined parts. This is much more efficient than modeling all of the geometry explicitly.

VRML models using the CIS2 related VRML prototypes represent a steel structure much more efficiently, effectively, and intelligently than the VRML models exported from a CAD program. Therefore much larger models can be displayed in a VRML browser on a handheld computer. With this capability, 3D visualizations of CAD models of steel structures can be viewed in the field at a construction site.

5. EXAMPLES AND CONCLUSIONS

The following VRML models were displayed on a Compaq iPAQ H3670 PocketPC [27] with 64 MB of memory running Pocket Cortona. The memory allocation was adjusted so that 40 MB was available to run programs. All steel parts with the same cross section have the same color, but all parts with the same color do not necessarily have the same cross section. All I- and T-beams are green, yellow, and cyan. All other sections are red, magenta, or blue except plates, which are gray.

Figure 1 is a VRML model of a simple three story steel structure with diagonal bracing. This model is small enough that the user can easily interactively move around the structure. Figure 2 is a VRML model of the steel support structure that is part of the emissions control system (ECS) ductwork for the NIST Large Fire Research Facility [28]. This model contains over 1600 parts and is a good example of the limitations of displaying VRML on a PocketPC. To minimize the number of polygons that are displayed, the thickness of the webs and flanges have been ignored. Nevertheless, it takes about five minutes for the model to display and is so large that the user cannot interactively navigate through the model. The only way to move around in the model is with predefined viewpoints in the VRML model. Small details such as the handrail look incomplete because of the low screen resolution. Figure 3 shows more detail of the structure in Figure 2. A close up view of a connection is shown in Figure 4. The location of the head of several bolts is shown as a black cross.

These examples show how 3D visualizations of steel structures can be displayed on a PocketPC. The VRML models are generated directly from CIS2 files and use VRML prototypes and other features to make the VRML files smaller and more efficient to process. Although there are limitations to the size of a VRML model that can be effectively displayed on a PocketPC, the inevitable improvements in processing power and graphics accelerators for handheld computers will provide an improvement for 3D visualization in the future. Within several years, it should be practical to make 3D visualization on PDAs a commonplace tool at construction sites.

6. REFERENCES

1. Microsoft Mobile Devices PocketPC, http://www.pocketpc.com/
2. Palm, http://www.palm.com/
3. Saidi, K., Haas C., Balli N., "The Value of Handheld Computers in Construction", in the Proceedings of the 19th International Symposium on Automation and Robotics in Construction, Washington, DC, September 2002.
4. Ames A., Nadeau D., Moreland J., "VRML 2.0 Sourcebook (2nd edition)", John Wiley & Sons, Inc., 1997.
5. Web3D Repository, http://www.web3d.org/vrml/vrml.htm
6. Cosmo Player, http://www.ca.com/cosmo/
7. Cortona, http://www.parallelgraphics.com/products/cortona/
8. blaxxun Contact, http://www.blaxxun.com/
9. Pivoron Player, http://www.nexternet.com/
10. Pocket Cortona, http://www.parallelgraphics.com/products/cortonace/
11. ATI Imageon 100, http://www.ati.com/technology/hardware/imageon100/index.html
12. IBM Microdrive, http://www.storage.ibm.com/hdd/micro/index.htm
13. OQO Corporation, http://www.oqo.com/
14. Pena-Mora F., Dwivedi G.H., "Multiple Device Collaborative and Real Time Analysis System for Project Management in Civil Engineering", Journal of Computing in Civil Engineering, January 2002.
15. Arc Second PocketCAD, http://www.pocketcad.com/
16. GiveMePower PowerCAD, http://www.givemepower.com/
17. Fathammer Ltd. X-Forge Technology, http://www.fathammer.com/x-forge/
18. Superscape Swerve, http://www.superscape.com/
19. "Pervasive 3D Graphics", Visual and Geometric Computing Group, IBM T.J. Watson Research Center, http://www.research.ibm.com/vgc/pdaviewer/
20. ISO/IEC 14772-1, "The Virtual Reality Modeling Language (VRML)", 1997, http://www.web3d.org/vrml/
21. Visualization of Steel Structures from CIMsteel Integration Standards (CIS2) Files, http://cic.nist.gov/vrml/cis2.html
22. Lipman R., "Mapping VRML Prototypes to the CIMsteel Integration Standards", to appear.

23. Lipman R., Reed K., "Visualization of Structural Steel Product Models", Electronic Journal of Information Technology in Construction, to appear.
24. Crowley A., Watson A., "CIMsteel Integration Standards Release 2", Steel Construction Institute Publication P268, 2000, http://www.cis2.org/
25. Eastman, C., "Building Product Models: Computer Environments Supporting Design and Construction", CRC Press, 1999.
26. AISC Electronics Data Interchange, http://www.aisc.org/Content/NavigationMenu/Steel_Solutions_Center/Time,_Money_and_Management/Electronic_Data_Interchange/Electronic_Data_Interchange.htm
27. Compaq iPAQ, http://athome.compaq.com/showroom/static/iPAQ/handheld_jumppage.asp
28. NIST Large Fire Research Facility, http://www.bfrl.nist.gov/info/facilities.html#burn

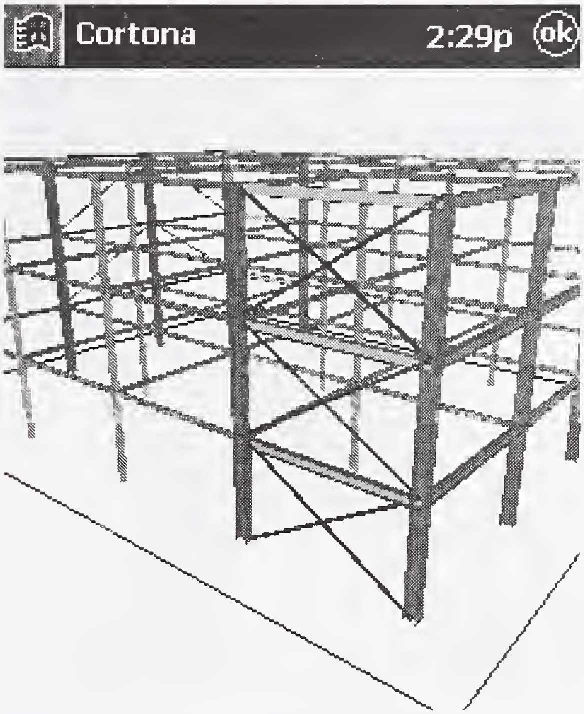


Figure 1. Three story structure with diagonal bracing.

Figure 2. Steel support structure for the NIST ECS.

Figure 3. Detail of the steel support structure for the NIST ECS.

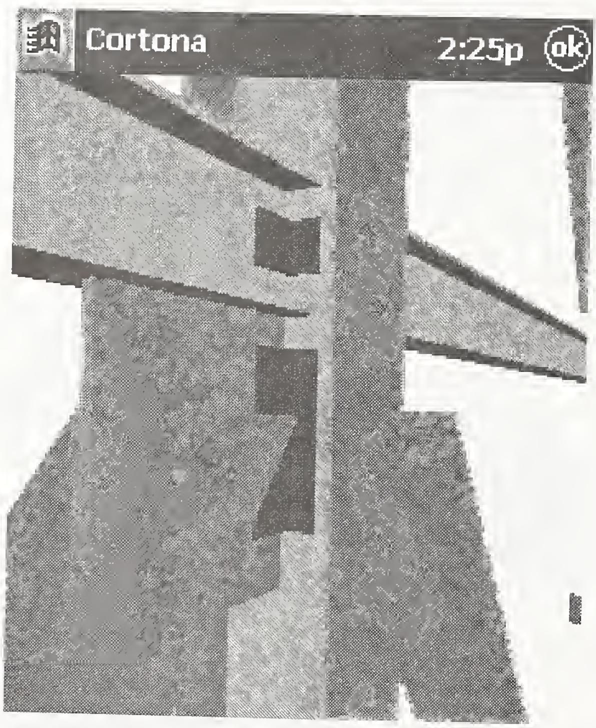


Figure 4. Close up view of a connection.

DIGITAL ARCHIVAL OF CONSTRUCTION PROJECT INFORMATION

By

DeWitt Latimer IV[1], Chris Hendrickson[2]

ABSTRACT

The Digital Library Initiative (DLI) concerns itself with the automatic creation, organization, and indexing of complex collections of data. With its own wealth of data types, the construction industry pushes the boundaries of the current technologies, especially with regards to automatic archival of construction project information. Many issues unique to the construction projects, such as data provenance and multi-media searching, require the integration of many technologies being developed at the various DLI funded research universities. This paper seeks to define those issues, the research being done, and provide some direction for future work in this area.

KEYWORDS

Digital library, document management, project document archival

1. INTRODUCTION

Construction projects generate volumes of information on a regular basis. However, access to this information can be difficult or awkward. Typically a person must go to a document area and rifle through papers, copies, or other documents to locate a specific document. Even worse, if a specific item is to be investigated, then the researcher must know where to search for information. If searching for information about a defect in the built entity, the researcher's information may be in the construction plans, meeting minutes between the contractor and the owner, in a bill of materials for parts relevant to the defect, inspector's reports, or even in a pre-construction environmental study. This problem is exasperated by the many media types common in the construction industry, application data files (such as CAD models, project schedule databases, or analysis calculations), images, structured formal documents, analysis results, and even audio or video. Although current web search technologies are becoming more sophisticated, these methods do not necessarily extend well to multiple distinct types of media. In a sense all these various media are the "documents" of the project.

This search problem is even an issue for construction document management. Current efforts in document management systems focus on 4 main issues: central storage of documents, routing of documents to interested parties, document version control, and security of documents. On the other hand, archiving is not only concerned with the saving of all versions of all documents over time but also indexing and providing an interface that allows for search across all the documents. Needless to say, the additional effort, and therefore cost, needed to properly archive documents is outside the budgets of construction projects. What is needed is a means of automatically generating archive documents quickly without much effort.

The digital library initiative [4] is sponsored by many federal agencies, including the National Science Foundation, DARPA, NASA, the National Library of Medicine, and the Library of Congress, among other federal agencies and private sector interests. Currently in its second phase, the goal of this program is to develop digital library technologies on a broad base to support the growing need for access to information. Many university research efforts are sponsored through this initiative. These universities are developing the base technologies and sciences to support the generation of these archives. Much attention is being paid to multi-media archives and archives with little, user-provided structure.

The remainder of this paper is organized in 4 sections. Section 2 provides the goals for archiving the documentation of a construction project as well as a taxonomy of the types of documents for typical construction projects; addressing these goals are the requirements for successful implementation of a construction digital library. Section 3 briefly outlines some commercial systems to indicate where these systems fail to meet the goals. Section 4 outlines some of the current research being performed that may be applicable to solving the issues in archival of

[1] dl4s@andrew.cmu.edu
[2] cth@andrew.cmu.edu
Department of Civil and Environmental Engineering
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh PA 15213

construction projects. Section 5 closes the document by stating some conclusions from the material presented and suggests directions for future work.

2. ARCHIVING A CONSTRUCTION PROJECT

2.1 Goal

The goal of a construction project archive is to provide a searchable data repository that represents both the final constructed artifact and the process that was used to construct the artifact. This archive must be constructed from the materials provided and yet must minimize the user's involvement in the creation of indices or other organizational tools (including keywords) needed to operate the archive. Care must be taken to ensure that the archive can handle the wide scope of items to be covered as well as able to manipulate the digital representation of those items (i.e. the various media formats).

2.2 Construction Information

Table 1 indicates a list of items that can be expected on a typical construction site. This table was inspired by the work of de la Garza and Howitt [18].

Request for information	Materials Management	Equipment Management	Cost Management	Submittals	Safety	QC/QA	Interpersonal	Schedule means and methods	Jobsite record keeping	Future Trends
Design intent and clarification	Access to material management	Equipment location	Budget	Test results	Accident reporting	Initiate inspections	Emails	Schedule	Recording timesheets	Positioning data
Subcontractor information	Material location	Fuel monitoring	Material cost accounting	Revisions to submittals	Reporting violations	Report QC/QA problems	Voice Mails	Schedule updates	Progress reporting	Sensory data
Contract specifications	Material order status		Equipment cost accounting	Physical Artifacts	Safety Plans	Reporting inspections results	Meeting Minutes	Delay recordings	Exception reporting	
Contract drawings	Request materials to site		Personnel cost accounting	CAD Models		Test Artifacts		As-built records	Visitor's log	
Work package information	Place material orders		Purchase Orders			Test Plans		Productivity information	Daily Construction Reports	
Means and methods	Material Specifications					QC/QA Plans			Images / Photos	
Implementation problems										

Table 1 – Typical documents of a construction project, by type

From the construction documents listed in Table 1, a list of digital data/file types was generated [Table 2]. File types are the digital representation of documents, as they would be stored in a digital archive. It is essential for any archival system to cover the breadth of these file types, given that each may contain information that is only referenced in that specific type of file.

Text	Digital Images	Application Data Files	Voice	Video	Formal Structured Documents
ASCII	BMP	MS-Project Files	WAV	MPEG	XML
MS-Word	JPG	VRML	MP3	AVI	CIM-Steel
PDF	TIFF	Pro-E	Real Audio	Real Video	spread sheets
Word Perfect	GIF	AutoCAD	AU	MOV	CIS/2 documents
...	...	...	...	...	...

Table 2 – Digital file types by application

A system must eliminate or mitigate all the fallacies, handle all the formats and document types would be a necessary requirement/criteria for successful implementation of a construction digital library.

3. COMMERCIAL SOLUTIONS

Several commercial products exist that provide document management services, but do not pass muster as an automated archival tool. Yet these products are likely to evolve over time, as the research described in section 4 matures.

JobDocs.com [8] provides an image management system for clients. While this service provides an engine for storing digital imagery about a job site, it does not automatically link the images with other project information, such as CAD models, product information, or meeting minutes. Additional storage of image versions is apparently handled by the user and not provided by the system.

Autodesk [1] provides a project collaboration environment for the management of print materials. While the environment does provide for managing document versions, the service does not provide for searching the contents of the documents themselves.

Meridian project systems [9] provides products for project management. The ProjectTalk [13] suite offers document management [14] functions. However, like Autodesk, this system is limited in its approach to searching for information. ProjectTalk behaves more like a file repository than a searchable archive.

In addition to the above systems, Amor/Faraj [17] provides a look into some of the reasons that integrated project databases (document management systems) fail. After some revision and simplification, we can identify 10 fallacies that may lead to a failure of integrated project databases:

1. An object-oriented system provides the complete solution;
2. A coordinated model solves representation problems;
3. A complete model of reality required;
4. User views are reconcilable;
5. The Internet solves the communication problem;
6. Printed documents will disappear;
7. CAD is the center of the archive;
8. No data ownership problems exist;
9. A database guarantees coordinated, consistent information;
10. The industry is ready to adopt such databases.

We will address these points in turn.

An object-oriented system provides the complete solution - From a computer science perspective, object-oriented system development is a method to develop software systems, and in itself does not solve any problems.

A coordinated model solves representation problems - Coordinated model(s) for very complex systems are difficult to create as can be evidenced by the slow development of industry foundation classes (IFC) [20] and the standard for the exchange of product data (STEP) [19] product descriptions; plus these representations, and systems based on these representations, will need constant updating as building technologies change.

A complete model of reality require - Completely modeling the reality of the site is not likely to be possible, however, any candidate system should be able to operate in a partial-information world.

User views are reconcilable - User views may not be reconcilable (certainly unless there is an information model).

The Internet solves the communication problem - The Internet does not, by itself, solve data communication issues, but is a tool that can be used to build systems that solve the problem.

Printed documents will disappear - Paper documents will likely still exist even with an easy access digital system.

CAD is the center of the archive - While useful for laying out spatial problems, CAD also cannot serve as a central data model for items such as equal opportunity reports, administrative budget items, and other items may not be linkable directly to structures.

No data ownership problems exist - In the business world, not all data may be viewable to all users, as the data may be proprietary, sensitive, confidential, or otherwise restricted for legal or business reasons.

A database guarantees coordinated, consistent information - Databases do not inherently solve the data consistency issue; by themselves databases do not guarantee the consistency and coordination of data, but respond to the data models with which they are designed.

The industry is ready to adopt such databases - Finally, the construction industry is notoriously IT poor; end users may not be receptive or have the inherent computing capability to support and use these IT systems.

4. DIGITAL LIBRARY TECHNOLOGIES

Many universities are participating in the Digital Library Initiative, however research at seven universities (Carnegie Mellon, Columbia, Cornell, Penn, Stanford, Berkeley, and UC-Santa Barbara) is quite relevant to the development of a digital archive of construction projects. These projects were noted because they directly answered one or more of the challenges presented in the previous section.

4.1 Carnegie Mellon University

The Informedia-II Digital Video Library [6] project at Carnegie Mellon University (CMU) focuses on building digital libraries from video sources. The project is automatically indexing and providing searchable interfaces to video news media. These indices are built by segmenting the video stream and using speech recognition to automatically generate transcripts of the sessions. In addition to normal textual querying of the archive, visual interfaces are being developed to query the archive with an image as the key. This visual querying works by having the user selecting a figure of interest in the video stream (typically a face) and then asking for matches. This provides a multi-media querying interface.

In a related project at CMU CCRHE (Capturing, Coordinating, and Remembering Human Experiences) [7], the video stream is augmented with sensory and position data. In this case, querying for video can be done by querying for the geographic position. As intended, this would be for looking for a memory of a trip; however in a construction application this capability could be used to ask for video from the location of a specific building feature.

4.2 Columbia University

PERSIVAL (PErsonalized Retrieval and Summarization of Image, Video And Language resources) is a patient care digital library [10] being developed at Columbia. This library take a pre-existing user model of health care professionals and provides an interface that attempts to predict the user's needs and interests. The information presented is from an online collection of patient records maintained by the hospital.

These records, in addition to information brought in from journal and web pages, are on distributed systems covering many different media (voice, image, text, etc.). Attention must be given to fusing together potentially conflicting information; methods such as automatically identifying source type, quality, and level of intended audience is needed so that the system can present information appropriate for the type of user performing the query (from the user model).

Other component technology to be developed are multi-modal query input, automatically augmenting queries based on current patient records, search and presentation of multi-media resources, supporting browsing over automatically constructed categories, graphical layout of multi-modal material, presentation and summarization of multiple relevant documents, merging repetitive information, and generation of appropriate level summaries.

The user based model approach may be helpful in the construction context since different participants have very different interests.

4.3 Cornell University

The Prism project [12] at Cornell focuses on digital library research in many areas. Some areas that are most relevant to construction project archiving are digital object architecture, human-centered research, and policy expression and enforcement.

The FEDORA (Flexible and Extensible Digital Object and Repository Architecture) project focuses on reliable and secure means to store and access digital content. This project seeks to address the creation of objects that aggregate heterogeneous types of data from distributed sources. Then the architecture allows objects to have global and/or domain-specific behaviors. The architecture is designed to support multiple viewing modalities, based on client access. The architecture utilizes a rights management model to aid in the creation of a "view of an object" contents.

The human-centered research group has deployed a field access system that utilizes global position sensing (GPS) and laser tracking (for indoor applications) to provide context-aware delivery of resources. On a construction site, this may be used to queue the plans to the room in which the person is currently located. Further, the group has studied the search patterns of users based on resource availability, relative expertise, and user characteristics (such as gender), which can greatly help deliver the right information in the field to individuals.

The policy expression and enforcement group focuses on a suite of tasks to deliver automated policy enforcement for digital libraries. The first task is generating a formal definitions and declarations, which provides a policy specification. Then, automated enforcement agents must be capable of understanding and acting upon the declarations. These behaviors manifest by the system allowing and encouraging permitted actions and denying unauthorized actions. The exact details of this behavior can be complex when considering distributed, mobile objects with various methods available dependent on the security access of the user. This work is moving towards a formal logic for such security specification.

4.4 University of Pennsylvania

The PENN Database Research Group has undertaken the exploration of data provenance [11]. Data provenance is the path the data took to become in the form being viewed. The provenance answers questions such as where and how the information was produced, who has corrected it, and how old it is. This tracking is essential to understand the past history of an item in the database, as well as fully appreciate and rely on the current state of the item.

4.5 Stanford University

The Stanford Digital Libraries Project [15] is looking into issues of interoperability, mobile access to digital libraries, and archival.

Interoperability thrust is focused on the SDLIP (Stanford Digital Library Interoperability protocol). SDLIP describes an interface for clients to request searches be performed over multiple clients by a library service proxy. This proxy would then communicate, using protocols defined elsewhere (such as Cornell's FEDORA[12]) to contact the data sources and perform the actual query to the databases. This protocol is already past its first release and is being refined through involvement with other, non-Stanford libraries.

The mobile access to digital libraries thrust is spearheaded by the development of a power browser for personal data assistants (PDAs). This power browser integrates automatic generation of indices of web pages with the necessarily limited resources of remote PDA computing.

The archiving thrust is concerned with the development of methods for local and wide-area archiving of file systems and replication of archives. Further, this thrust is concerned with the generation of cost/benefit analysis tools, such as a simulator, to help guide the design of archives.

4.6 University of California at Berkeley

The Digital Library Project at Berkeley [16] has three research areas of interest: computer vision in digital libraries [3] and a geodetic information system (GIS) viewer [5].

The use of computer vision in creating a digital library enables the creation of large, automatically annotated data stores from digital images. Performing object recognition on images enables each image to be labeled with its contents. From here, interfaces for statistical based queries on the data are being developed. Further work is being done to be able to query a database for images relevant to a passage of text (such as one from another document).

The GIS Viewer 4.0 is designed as a tool for displaying, manipulating, editing, querying, and otherwise interacting with layered geo-spatial information. In construction, spatial information is normally presented in various layers, floors, electrical, plumbing, etc.

4.6 University of California at Santa Barbara

The Alexandria Digital Earth Prototype [2] project at UC Santa Barbara is seeking to develop a digital library environment and service based on the digital Earth metaphor. The focus of this research is to create and use Iscapes (Information Landscapes), which are a personalized digital information collection. These Iscapes are intended to have layered services which organize, access, and use different types of information. Currently, the project is focusing on creating these Iscapes for learning/educational purposes in undergraduate classes.

5. CONCLUSIONS

From the survey of digital library technologies, there is much support for the argument that a digital archive, or library, can be built for construction projects. Although research is not currently directed at this specific domain, the research of the many universities provide the key technologies to solve most of the 10 issues in developing an integrated project database brought forth by previously [Table 3].

Issues	Technologies	Universities
1. OO does not provide complete solution	Developing paradigm independent models of archives	Cornell
2. Coordinated model does not solve representation	Do not assume documents conform to a single structure	Cornell, Stanford
3. Complete model of reality required	Create reports and views based on partial information	Columbia
4. User views are reconcilable	Architectures to support multiple user views	Columbia, Cornell, Berkeley, Santa Barbara
5. Internet solves communication problem	Specific digital library/archive protocols for communication	Stanford
6. Paper documents will disappear	Multi-media archiving	CMU, Berkeley, Santa Barbara
7. CAD is the center of the database	Loosely distributed document structure	Cornell, Stanford
8. No data ownership problems	Data security, migration, and ownership model	Cornell
9. Database guarantees coordinated, consistent information	Data provenance theories to aid in maintaining data	U Penn
10. Industry is ready to adopt such a system	Not Addressed	NA

Table 3 – Technologies to Address Issues of Integrated Project Databases

Although object oriented systems are heavily promoted to be the principal architecture upon which a library is built, there is no assumption of the structure of the underlying data components. Indeed, the systems that propose object structures for the digital library interaction assume no structure for the underlying objects, which are normally items such as web pages, digital books, or image files.

While all digital library systems propose a model of the library, almost all research into searching the library and structuring the library is by viewing the underlying documents as "unhelpful." In this case, unhelpful means that the document contains little or no meta-information specifically designed to help a search engine, or to help build links between different documents. On a higher level, this means that no one model of underlying documents is assumed.

Some research is being done to address answering queries when incomplete information is available. Specifically the work at Columbia targets generation of reports based on potentially incomplete information. Also, the work at Penn on data provenance gives client applications the ability to determine the sources, and thus the potential gaps, in the knowledge available.

In almost all systems, a single user view is not assumed, and a great deal of attention is being devoted to the creation of personalized user views. This will allow disparate user views to be built from various user profiles to suit the needs of the various users.

The current architecture research being performed in digital libraries understands that the internet is not a panacea that solves all networking problems. Indeed the current architecture projects appear to be very cognizant of the need to develop reasonable protocols for interacting with the digital libraries.

Given the highly distributed nature of the digital libraries being considered, no one data item appears to be driving the structuring of the archives. Thus, there is no misconception that CAD, or any other document type will be central to the archive.

The data security work at Cornell is specifically targeting the creation of ownership and access management technologies. With these technologies, a system can be envisioned which will allow contractors, subcontractors, architects, engineers, owners, and other interested parties to be able to connect systems together with each managing their own access rights. And upon commissioning, the mobile library projects at several universities can make sure that information allowed to be mobile is automatically migrated to the owner's systems.

Penn's work on data provenance directly addresses the coordinated and consistent information issue. When the provenance work is combined with the Columbia work, which focuses on ranking various input sources for appropriateness and handles inconsistencies among sources, then there is the ability to generate a system that can recognize and handle coordinating information to present consistent information to the users.

Two areas that are lacking in the current research are very specific to the construction industry and must be solved in the scope of developing a library specifically for construction projects: the issue of physical documents and the readiness of industry to adopt such a technology. In order to successfully manage documents and ensure they are entered into the system, a series of procedures and field protocols need to be developed. To prepare the industry to receive this research will require economic impact analyses that are outside the scope of the current research.

Overall, there is clearly a critical mass of technology being developed that can support the creation of a digital archive for construction projects. The next step should be the formalization of requirements for such an application and the development of a prototype based on the components already available and/or under development.

6. REFERENCES

[1] "Autodesk BCS: Project Collaboration and reprographic print management for building, design and management for building, design, and management professionals", http://www.buzzsaw.com/, accessed on 2002-03-30

[2] "Award #9817432 – DLI Phase2: Alexandria Digital Earth Prototype", https://www.fastlane.nsf.gov/servlet/showaward?award=9817432, accessed 2002-03-30

[3] "Computer Vision Meets Digital Libraries", http://elib.cs.berkeley.edu/vision.html, accessed 2002-03-30

[4] "Digital Libraries Initiative Phase 2", http://www.dli2.nsf.gov/, accessed 2002-03-30

[5] "GIS Viewer 4.0", http://elib.cs.berkeley.edu/gis/index.html, accessed 2002-03-30

[6] "Informedia Digital Video Library at CMU", http://www.informedia.cs.cmu.edu/, accessed 2002-03-30

[7] "Informedia-II Digital Video Library", sub-title: "CCRHE – Capturing, Coordinating and Remembering Human Experience", http://www.informedia.cs.cmu.edu/ccrhe/index.html, accessed 2002-03-30

[8] "JobDocs.com online construction documentation image management system", http://www.jobdocs.com/, accessed 2002-03-30

[9] "Meridian Project Systems - Comprehensive solutions for organizations that rely on successful projects", http://mps.com/, accessed 2002-03-30

[10] "PERSIVAL – A Patient Care Digital Library", http://www.cs.columbia.edu/diglib/PERSIVAL/, accessed 2002-03-30

[11] "Penn Database Research Group: Data Provenance", http://db.cis.upenn.edu/Research/provenance.html, accessed 2002-03-30

[12] "Prism Home Page", http://www.prism.cornell.edu/, accessed 2002-03-30

[13] "ProjectTalk.com – where the project team meets", http://www.projecttalk.com, accessed 2002-03-30

[14] "ProjectTalk.com | Project Management | Document Management", http://www.projecttalk.com/MK/pm/doc.shtml, accessed 2002-03-30

[15] "Stanford Digital Libraries Project, The", http://www-diglib.stanford.edu/, accessed 2002-03-30

[16] "UC Berkeley Digital Library Project", http://elib.cs.berkeley.edu/, accessed 2002-03-30

[17] Armor, Faraj, "Misconceptions About Integrated Project Databases", Electronic Journal of Information Technology in Construction, Volume 6, 2001, http://www.itcon.org/2001/5/paper.pdf, accessed 2002-03-30

[18] Garza, Howitt, "Wireless communication and computing at the construction jobsite", Automation in Construction, Volume 7, Number 4, pg. 327-347

[19] "Industrial automation systems and integration - Product data representation and exchange", International Organization for Standardization, Standard 10303-1:1994

[20] "IFC Object Model Release 2.0", Industry Alliance for Interoperability, CD-ROM, 1999

E-WORK: THE NEXT ITERATION IN CONSTRUCTION MATERIALS MANAGEMENT SYSTEMS

Daniel Castro-Lacouture[1] and Miroslaw J. Skibniewski[2]

[1] *Ph.D. Candidate, School of Civil Engineering, Purdue University, West Lafayette, Indiana, 47907-1294, USA*
http://bridge.ecn.purdue.edu/~castrola

[2] *Professor of Civil Engineering, Purdue University, West Lafayette, Indiana 47907-1294, USA*
http://bridge.ecn.purdue.edu/~mirek

Abstract: The initiative of implementing Internet-based solutions for the integration of the supply chain of materials will not succeed if there is not assurance of collaboration and communication among the parties in the construction business. Neither EDI nor web-based supply chain applications can facilitate both the process integration and flexibility required by the construction processes practiced today. E-work is a suitable concept for this endeavor because it comprises principles that allow cooperation and collaboration in the organization. The research described in this paper follows upon recent publications on E-work models in order to propose a set of applications in construction materials management systems. Expected benefits are related to conflict resolution and reduction of lead times and subcontractor overhead costs. Implications of this work with respect to the initiatives based on IAI/IFC or STEP models comprise the potential for prediction of materials requirements with the use of task administration protocols and autonomous agents.

Keywords: E-work, materials management, protocols, agents

1. INTRODUCTION

1.1. Supply Chain of Materials in Construction Processes

Construction processes rely in a great extent in the exchange of information and permanent interactions of entities and resources. Even though the configuration and set up of machines is usually different than in the manufacturing industry, the interrelations of relevant participants in a construction process are considerable, and their management will have a direct impact on the success of the project in terms of time, cost, quality and morale.

There are other considerations that need to be addressed when designing a construction materials automation system, such as the effect of design in the construction process, the isolation of materials procurement from design, differences in CAD formats, use of non-integrated databases, brief and usually poorly coordinated contribution of subcontractors in the design and procurement processes, etc [3]. These problems are well documented in the construction literature, and have been in recent years a subject of study of several business organizations such as the Construction Industry Institute in the United States and related bodies worldwide. The initiative of implementing a collaborative e-business solution for the integration of the supply chain of materials, such as steel reinforcement rebars, will not succeed if there is no assurance of collaboration and communication among the players in the construction business [3].

It is possible to deploy an e-commence function with the objective of performing transactions between contractors, suppliers and clients. In fact, that can be done easily and is currently being utilized by construction material suppliers and contractors. Ordering, billing and information-sharing functions are common in the form of Web-based applications [11]. However, by merely setting up a transaction platform framed by e-commerce principles, the success of the integration of critical stages in the construction process is not necessarily guaranteed. Moreover, it is needed to consider a new concept that fosters effectiveness in the overall supply chain system integration, that is, to achieve adaptability of resources throughout the supply chain of materials in the overall construction process, or to rapidly adapt to changes in the required product quantity and design [10]. This concept must integrate the general and specific project design in order to avoid conflicts in the subsequent stages of the project development.

The objectives of the materials supply chain automation consist of minimizing processing costs, enhancing delivery and response times and improving communication in order to expedite conflict resolution. Although the scope of this research is confined to describe the use of E-work models, metrics associated to these model objectives will be originated from time studies, workflow simulation, activity-based costing and frequency of documented delays due to conflicts. The idea of automating the supply chain of materials must consider the multiple sources of disagreement related to negotiation protocols among contractor, subcontractor, material supplier, designer and client, as a consequence of the struggle to obtain better prices and product quality. An electronic market system can reduce customer's cost of obtaining information about prices and product offerings of alternative suppliers, as well as supplier's costs of communicating information and negotiating about the prices and product characteristics [2].

1.2. E-work and Autonomous Systems

E-work is composed of the collaborative, computer-supported activities and communications-supported operations in highly distributed organizations, and also investigates fundamental design principles for their effectiveness [13].

The goal of E-work for this research is to integrate the components of the construction supply chain of materials, from structural design to procurement and assembly. The utilization of this approach is justified by the assumption that construction processes are composed of highly distributed organizations of autonomous systems. This assumption is based on the unique nature of the construction industry in terms of interactions among business partners, clashing interests, isolated design and procurement, etc. In addition, every construction project is different from another. The materials, procedures, contractors, design firms, or contracts in construction projects may be similar, but the final product is evidently different. It is a customized product that has to evolve from the initial stages of client briefing to the complex states of seismic resistance, fit outs and finishes. That is one strong reason to consider this organization as complex and composed of autonomous systems. These systems are composed of the equipment, hardware or software utilized by various designers, suppliers, subcontractors or contractors in order to achieve the goal of project completion [3].

Figure 1 shows a sample of the numerous interactions among project stakeholders in order to approach the process of design, revision and procurement of materials for the structural reinforcement division of a standard construction project.

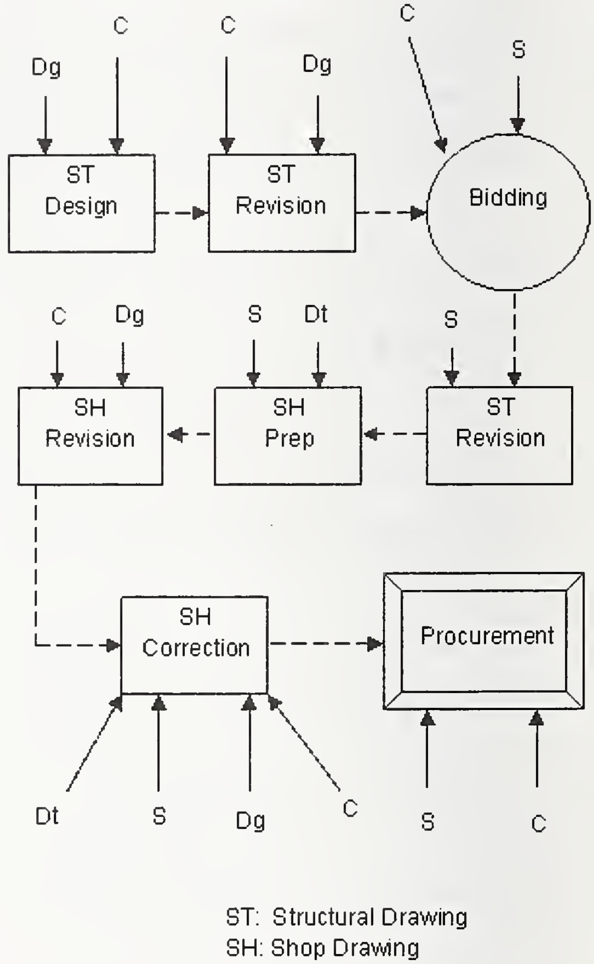


Figure 1. Scheme of interactions during the design, revision and procurement of materials on a standard construction project.

The terms C, S, Dg and Dt stand for Contractor, Rebar Supplier, Designer and Detailer.

2. MODELS OF E-WORK

2.1. Task Administration Protocol (TAP)

Task Administration protocols are defined as the logical rules for the workflow control that enable effective collaboration by communication and resource allocation among production tasks [14]. Effective collaboration, in the context of a supply chain, stands for obtaining real-time information, minimizing cost, increasing levels of service, improving communication and enhancing delivery and response times [8]. Within the context of the construction process and in particular for the case of construction materials management, the following protocols are defined in Table 1.

Table 1. E-work solution algorithm with Task Administration Protocol [6].

Types	Protocols
Inquiry	A → B: request X B → A: answer (X) A,B - agents X – questions
Passing	A → B: tell X A,B - agents X – message
Negotiating	(1) A → B, C: tell X (2) B → A: If ValueB(X) OK, Tell 'acceptB' else Tell X' (3) C → A: If ValueC(X) OK, Tell 'acceptC' else Tell X'' (4) A → B, C: If acceptB and acceptC, end. If acceptB and X'', X=X'', go to (1) If acceptC and X', X=X', go to (1) If X' and X'', X=(X',X''), go to (1) A,B,C – agents X,X', X'' – alternatives
Announcing	A → B,C,D: tell X A,B,C,D - agents X – message
Approving	(1) A → B: request X (2) B → A: If ValueB(X) OK, tell 'approveB' else X' (3) A → B: If approveB, end. If X=X' go to (1) A,B – agents X,X' – alternatives, where X' is lacking approval

In order to improve performance, the protocol can trigger and initiate necessary and timely interaction tasks under coordination logic. In order to explain and illustrate an E-work solution with task administration protocol (TAP) it is important to clarify the concept of agent. An agent is a 'program' (autonomous entity) that represents a group of parties [6]. Parties may include 'resources', e.g. humans, machines and computer systems. The role of the agents is to cooperate with each other through protocols to perform specific tasks and to achieve the system's goal. Along the lines of this definition, the active tasks administration protocols (TAP) control rules that tenable effective collaboration among agents and tasks. Protocols provide the rules for agents to procure information [6]. A typical scenario of material procurement, illustrated in Figure 2, may involve a contractor and several steel suppliers. The contractor inquires two rebar supplier agents (a_1 and a_2) about quantity take-off of some rebar material. The two supplier agents respond the request. Based on the responses, a_0 does some checking and sends the decisions and approvals to the three supplier agents.

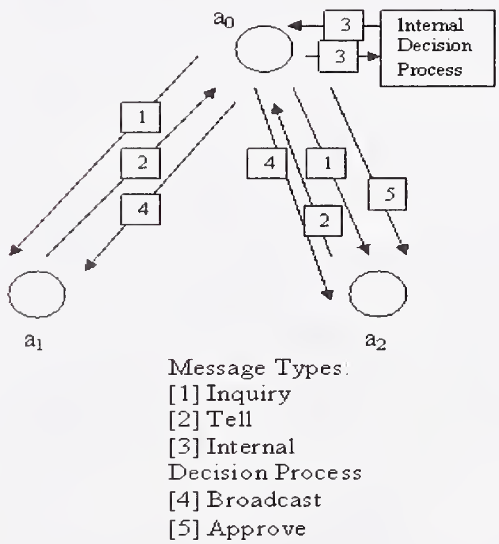


Figure 2. Task administration protocol applied to construction materials management automation.

2.2. Multi-agent Model

Agents are defined as a hardware or software system that can act, interact and respond according to given goals [14]. For the particular case of construction materials management networks, the agents need to feature autonomy and flexibility. Software agents developed specifically for the supply chain may support the ability to closely integrate buyer and seller processes as well as to provide sufficient flexibility to keep up with changes to supply chain operations [12]. As a requirement to be autonomous, the agents need to have [6]: database and protocols, sufficient capability for anticipating their long-term viability domain, measurable output (processing time, waiting time, etc.), ability to learn and a sensorimotor apparatus (protocols, databases and human observation can serve to this purpose). A similar case takes place in construction processes, where the final decision on material quantities is not final until the material suppliers (for the case of steel reinforcement rebars) provide the estimation based on the structural design drawings. Therefore the system of materials procurement has to be initiated without the benefit of planning, which is what happens in most of the cases. An architecture consisting of four agents will be used to illustrate an E-work model for the construction materials management automation on the Internet. Agents to be considered are a forecaster agent, a recorder agent, a planner agent and a reviser agent. The flow diagram for these interactions is shown in Figure 3.

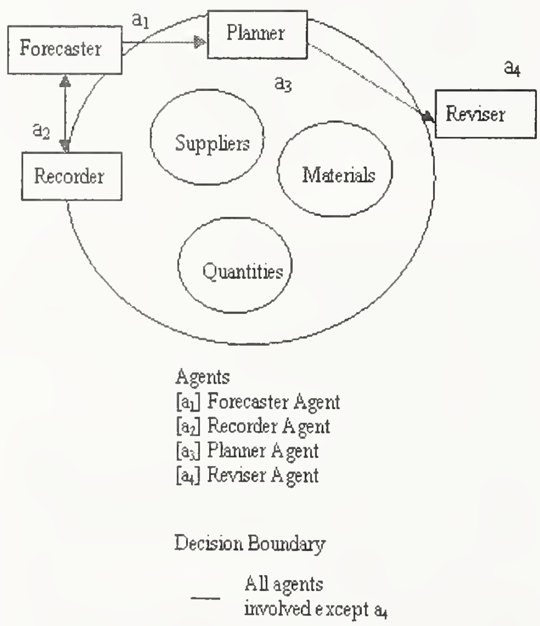


Figure 3. Multi-agent model applied to construction materials management

However, an E-work model of multi-agents can help predict the requirements of materials using knowledge of the supplier past estimations for standard or special cases, in addition to recent observations of other projects in progress or completed. These predicted job requirements can then be used for better planning [14]. In an effort to find the solution to a problem in a flexible manufacturing system, Esfarjani and Nof [4] found out that the challenge consists of managing the distributed material and information flow. The construction process, due to its diverse and almost unlimited source of ideas and methods for materials procurement, handling and procurement, qualifies to be treated as a flexible process.

2.3. Parallelism Model

An E-work model of parallelism to ensure effectiveness in the Internet deployment of a construction material management automation system must have into account that there are several e-activities present. It is possible to obtain improvements in time and accuracy through the utilization of a client-server model and the assignment of tasks to the available resources, i.e. material suppliers, designers, subcontractors or delivery entities. There will be a coordination agent that allows the parallel execution of tasks, thus having earlier progress or even completion dates. Parallel processing in supply chains has an explicit focus on collaborative problem solving between multiple processors. This problem solving is usually focused more on numerical processing than intelligent reasoning [12]. Figure 4 displays the structure of a parallel tasking, using master-slave architecture.

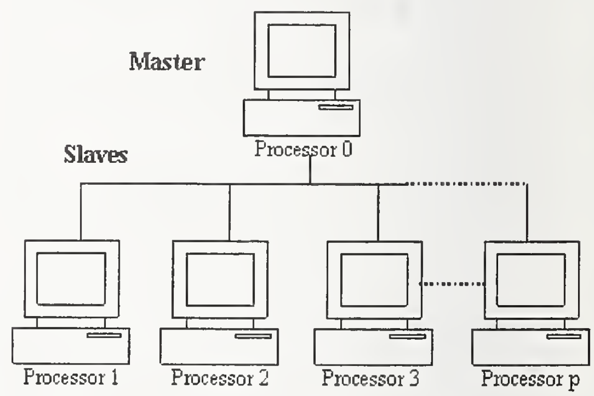


Figure 4. Master-Slave architecture in parallel multitasking.

In order to communicate and cooperate in the Internet, certain procedures must be established to coordinate the interaction of participants in order to effectively accomplish the tasks [16]. In this case there are various types of inquiry, approval or negotiation protocols that are presently used.

There is a need for a coordination protocol that will guide and support the interaction of participants in order to successfully perform the necessary tasks. The ultimate goal of the coordination-related research is to find the most effective coordination scheme that maximizes collaboration effectiveness among participants. The effectiveness in the collaboration among material suppliers, subcontractors, designers and contractors will have a dramatic impact in the overall success of the construction project.

The interaction diagram of a typical workflow for construction materials management shows the different interactions among resources. Figure 5 displays the interaction diagram for a typical supply chain of materials.

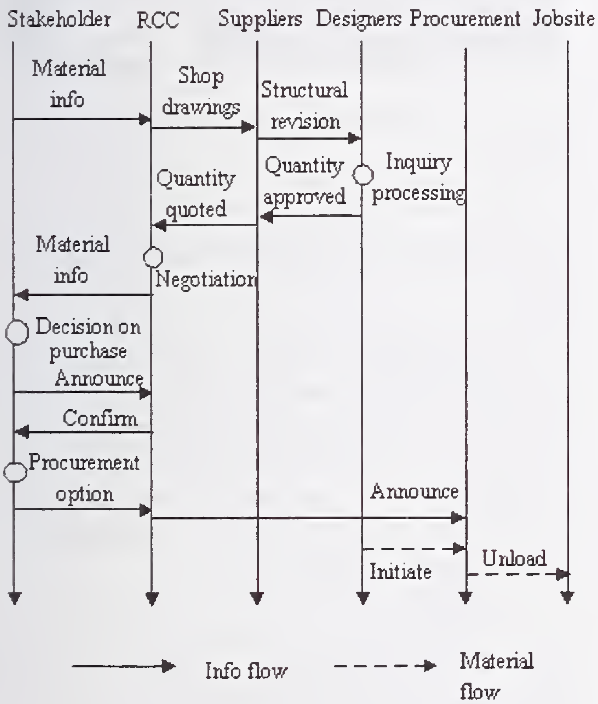


Figure 5. Interaction diagram for a typical supply chain of construction materials.

Interaction diagrams are tools that characterize the elements in the supply chain of information and materials. These tools may be utilized when deploying an e-business solution for the integration of the supply chain. For instance, any request by the contractor of information regarding quantity take-off from shop drawings will flow from the material supplier to the designer for revision and approval. Subsequently, the designer will send the approved drawings back to the supplier for delivery through a web operation center. This e-business approach is aiming to integrate the process of design, estimation, revision and procurement of materials from the early stages of the construction project. The interaction diagram also indicates the decision nodes in the overall process, i.e. inquiry processing, negotiation, decision on purchase and procurement option.

The research developed by the PRISM Lab at Purdue University [14] uses parallel computing to simulate autonomous agents. This tool, called Teamwork Integration Evaluation (TIE), focuses on measuring the coordination performance of all participants [1]. Figure 6 shows an adaptation of the original TIE/Protocol, with outputs relevant to the new application in the construction material automation process.

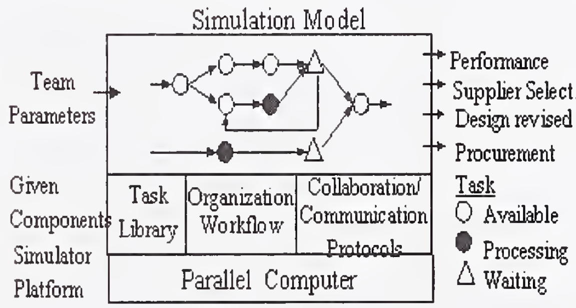


Figure 6. TIE/Protocol for construction material automation (adapted from Anussornnitisarn and Nof [1]).

2.4. Conflict Resolution Model

The inherent nature of construction processes plagued with conflicts, create an area for improvement. An E-work model of conflict resolution will be of great benefit for the automation of construction materials management systems deployed on the Internet.

A collaborative method for facility design is needed in order to increase design quality and shorten the design process development [14]. The process of design, estimation, revision and procurement of construction materials displays lots of conflict situations, which result from the interaction of cooperating designers, suppliers, subcontractors or contractors. The new method of conflict resolution, MCR, developed by Lara [9] and Lara et al. [10] comprises rational execution of pre-ordered conflict resolution approaches: direct negotiation, third party mediation, incorporation of additional parties, persuasion and arbitration or settlement of claims.

The main cost-differentiating factors influencing choices in conflict resolution architectures are: intensity of processing (measured by task complexity), intensity of communication (measured by quantity of exchanged communication) and extent of communication (measured by the level of networking in the process) [5]. This study points out that previous analyses in the area have failed to consider management costs, which can be dissimilar across different IT architectures. As it can be seen in Figure 7, the costs for conflict resolution increase as stages escalate from a direct negotiation to arbitration.

Cost Progression

Stage 1: Direct negotiation

Stage 2: Third party mediation

Stage 3: Incorporation of additional parties

Stage 4: Persuasion

Stage 5: Arbitration

Figure 7. Cost of conflict resolution stipulated in MCR [9].

This multi-approach method facilitates computer-supported conflict resolution. Its performance is improved through the inclusion of principles for prevention of conflict perpetuation. Along the same line of logic, the implementation of computer-based learning increases the usefulness of the method. As expected, the integration of conflict detection and resolution results in an increased effectiveness of the facility design process. The facility design language (FDL) provides a uniform support framework for communication among facility designers in a distributed environment, thus allowing distributed users to develop, evaluate, reconcile and modify CAD system models of the facility. This language fits the needs of the construction process due to its strong dependency on CAD drawings for information exchange and quantity take-off revision.

Figure 8 shows the arbitration steps to be included in the construction material process in order to incorporate prevention, therefore the system will be continuously providing feedback to the parties in order to learn from previous experiences.

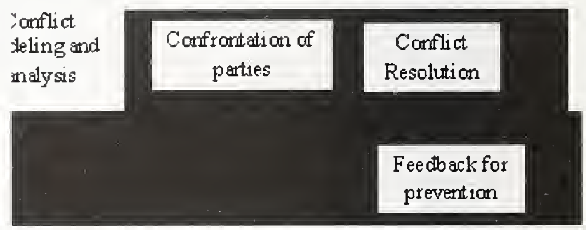


Figure 8. Sequence of steps in the prevention stage.

2.5. Model of Next Generation of ERP Systems

An E-work model that incorporates design considerations is desirable for the successful practice of construction materials management automation. As indicated by Nof and Anussornnitisarn [15], the key for the successful implementation of ERP systems is to be able to anticipate the problems.

For the particular case of construction materials management, some considerations need to be addressed in order to fit the purposes of the new generations of ERP systems. Firstly, when material suppliers are required to provide a take-off of a structural drawing, the design process should be able to detect, synthesize and prioritize the problems that might occur with the interacting participants. Secondly, the design should be able to provide the grounds for the protocols to be used and what and when to communicate. Priority schedules for the different participants of the construction process (contractor, subcontractor, supplier and designer) must be kept active. Third, the independent entities should be able to participate in the design process. This is an ambitious goal, but this way coordination with other agents will be possible, therefore problems can be anticipated and dealt with in a timely manner. Finally, the new ERP system will recognize conflicting intentions among distributed entities. If there is a pattern of miscalculations by a designer that is being rejected by the materials supplier, the ERP system should be able to trigger an action (agent) on the contractor, allowing the system to foresee an upcoming conflict therefore evaluate the design. A similar case may be encountered the other way around. If the material supplier is not providing an accurate take-off of materials, the ERP system triggers an action framed by an active protocol, anticipating the conflict and informing the supplier about the possible solution to the inconvenience. Figure 9 shows the concept of the next generation of ERP systems.

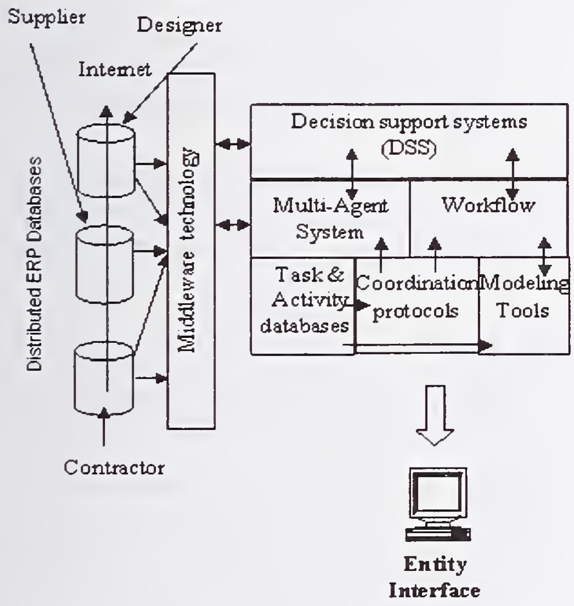


Figure 9. Next generation of ERP systems [15].

The communication infrastructure of parties involved in an integrated ERP system for the construction industry will depend on models such as the STEP, EDI or IAI/IFC. A subset of these models, SUMMIT, can be used to support the integration of supply chain management applications [7]. This approach developed a framework to automate the tendering, ordering, delivery, invoicing and payment processes of house systems, equipment and services in their supply chain, and where members have different backgrounds regarding information and communication technologies.

3. CONCLUSIONS

The supply chain of materials in construction processes remains a critical competency for the success of the construction business. Even though fast developments in communication infrastructure through models of EDI, STEP or IAI/IFC, the management and control of entities in the supply chain of materials are very often not carried out automatically. With the incursion of the Internet and its e-business capabilities, new approaches for the automation of supply chains of materials have been developed. However, the construction industry has yet to utilize intelligent tools for the resolution of common conflicts in the planning, design, procurement and delivery of materials.

E-work is a suitable concept for this endeavor because it comprises principles that allow cooperation and collaboration in the organization. E-work has provided benefits to the exchange of information and collaboration among autonomous systems in manufacturing organizations. The nature of construction processes in terms of occurrence of conflicts, autonomy of interacting systems, distributed organization and sequencing of tasks in the supply chain provides areas for applicability of E-work models. The research described in this paper builds upon recent work and publications on E-work models applied to systems of manufacturing organizations.

Autonomous agents can help predict the requirements of materials using knowledge of the supplier past estimations for standard or special cases, in addition to recent observations of other projects in progress or completed.

The sequencing of tasks for the exchange of information prior the procurement and delivery stages can be achieved via task administration protocols, whose defined steps convene more complex intentions than single communication acts.

Interaction diagrams provide information on different interactions among resources, allowing stakeholders to consider parallel processing for an effective automation of supply chain processes focusing on characterization of entities and decision nodes.

Finally, the process of design, estimation, revision and procurement of construction materials displays conflict situations that can be resolved with the implementation of multi-approach methods facilitated by computer-supported conflict resolution.

4. REFERENCES

[1] Anussornnitisarn, P. and Nof, S.: "Evaluator TIE/Protocol, a teamwork integration for coordination protocols", Proceedings of ICPR-2000, Bangkok, Thailand, August 2-4, 2000.

[2] Bichler, M. and Segev, A.: "Methodologies for the design of negotiation protocols on e-markets", Journal of Computer Networks, 37: pp. 137-152, 2001.

[3] Castro-Lacouture, D. and Skibniewski, M.: "Development of an e-business solution for the integration of steel reinforcement supply chain in construction projects", Proceedings of the First International Conference on Construction in the 21st Century (CITC2002), Miami, Florida, USA: pp. 197-204, 2002.

[4] Esfarjani, K., and Nof, S.: "Client-server model of integrated production facilities", International Journal of Production Research, 36(12): pp. 3295-3321, 1998.

[5] Francalanci, C., and Piuri, V.: "Designing information technology architectures: a cost-oriented technology", Journal of Information Technology, 14: pp. 181-192, 1999.

[6] Huang, C., and Nof, S.: "Formation of autonomous agent networks for manufacturing systems", International Journal of Production Research, 38(3): pp. 607-624, 2000.

[7] Jardim-Gonçalves, R., Tavares, R., Grilo, A., and Steiger-Garçao, A.: "A framework for integration of supply chain management applications based on the IAI/IFC model", Journal of Automation and Robotics in Construction, Vol 17: pp. 745-750, 2000.

[8] Lancioni, R., Smith, M., and Oliva, T.: "The role of the Internet in supply chain management", Industrial Marketing Management, 29: pp. 45-56, 2000.

[9] Lara, M.: "Conflict resolution in collaborative facility design", Ph.D Dissertation, School of Industrial Engineering, Purdue University, 1999.

[10] Lara, M., Witzerman, J., and Nof, S.: "Facility description language for integrating distributed design", International Journal of Production Research, 38(11): pp. 2471-2488, 2000.

[11] Lin, Y., and Tserng, H.: "A model of supply chain management for construction using information technology", 18th International Symposium on Automation and Robotics in Construction, Krakow, Poland, pp.141-146, 2001.

[12] Nissen, M.: "Supply chain process and agent design for e-commerce", Proceedings of the 33rd Hawaii International Conference on System Sciences: pp. 1-9, 2000.

[13] Nof, S.: "Intelligent collaborative agents", Research Memorandum, No. 99-8, March 1999, School of Industrial Engineering, Purdue University, 1999.

[14] Nof, S.: "Models of e-Work", Proceedings of the IFAC symposium on Manufacturing, Modeling, Management and Control, Rio, Greece, pp. 521-526, 2000.

[15] Nof, S., and Anussornnitisarn, P.: "E-Work: The challenge of the next generation ERP systems", Production Planning and Control, 2002.

[16] Nof, S.Y., and Ceroni, J.: "E-Work models and systems: Modeling manufacturing and service in the Internet Era", Elsevier (Forthcoming).

Capitalizing on Early Project Opportunities to Improve Facility Life-Cycle Performance

by

Calvin Kam[1], Martin Fischer[2], Reijo Hänninen[3], Seppo Lehto[4], Jarmo Laitinen[5]

ABSTRACT: This paper presents the findings from the design and construction of the Helsinki University of Technology Auditorium Hall 600 (HUT-600) in Finland. Running simultaneously with the construction project, an international research partnership extensively documented and analyzed the use of product modeling and interoperability standards for information exchange. The project team improved design cycle times and minimized data re-entry through an array of design, visualization, simulation, and analysis tools. Building on the resulting efficiency and time-savings during the early conceptual phase, the project team conducted a variety of in-depth life-cycle studies and alternative comparisons on thermal performance, operation costs, energy consumption, and environmental impact. Such unconventional practices empowered the building owners to better align the long-term facility values with their strategic plans.

KEYWORDS: Life-Cycle Analysis, Pilot Project, Product Model

1.0 INTRODUCTION

The design and construction of the Helsinki University of Technology Auditorium Hall 600 (HUT-600) adhered to the *Level of Influence on Project Costs* concept (Paulson 1976) that making a decision during the early project phase has a relatively high impact and low cost. This paper explains the Product Model and 4th Dimensional (PM4D) Approaches and PM4D Processes that were developed in the HUT-600 project. We document the life-cycle studies made during the early schematic design phase of HUT-600, highlight their relationships with PM4D Approaches and Processes, and analyze their implications on the long-term performance of the capital facility.

In autumn 2000, the Helsinki University of Technology requested its property owner, Senate Properties of Finland, to build a new multipurpose auditorium that would be capable of accommodating 600 people. Running on a fast-track design and construction schedule, Senate Properties assembled a team of designers, consultants, contractors, and researchers early during the conceptual planning phase. This partnership received financial supports from the Finnish National Technology Agency (TEKES), which sponsored the testing of state-of-the-art technologies and data standards in HUT-600. Since HUT-600 is the first institutional application of the Industry Foundation Classes (IFC) interoperability standards, it is also a construction pilot.

[1] Graduate Research Student, Civil & Environmental Engineering, Stanford University, Stanford, CA. 94305-4020, U.S.A. <ckam@stanford.edu>

[2] Associate Professor, Civil & Environmental Engineering and (by courtesy) Computer Science, Director of the Center for Integrated Facility Engineering, Stanford University, Stanford, CA. 94305-4020, U.S.A. <fischer@stanford.edu>

[3] General Manager, Olof Granlund Oy, Helsinki, Finland <reijo.hanninen@granlund.fi>

[4] Senate Properties, Helsinki, Finland <seppo.lehto@senaatti.com>

[5] Ph.D, Customer Solutions, Solibri Inc., Helsinki, Finland <jarmo.laitinen@solibri.com>

2.0 THE PM4D APPROACHES

The HUT-600 project team followed the PM4D Approaches in which they relied on product modeling and interoperability standard to eliminate the inefficiency and risks of data re-entry in conventional practices. The approaches aimed at leveraging upon state-of-the-art analytical and visualization tools to optimize the design, construction, and operation of the proposed facility during early project phases (Kam et al 2002).

The approaches recognized the importance of having cross-disciplinary expertise to improve design and construction integration. With the goal of improving the quality of design and construction services, the approaches called for utilizing intelligent object-oriented product models and the IFC interoperability standard that supported data sharing. Furthermore, the approaches adopted various visualization tools and analysis tools to additional life-cycle performance data to support the decision-making processes.

3.0 THE PM4D PROCESSES

The PM4D Processes explain the software applications, data standards, and information flow that took place in HUT-600. During the early schematic phase, the architects and mechanical engineers designed with object-oriented modeling software such as Graphisoft's ArchiCAD[6] and Progman Oy's MagiCAD[7]. IFC release 1.5.1 interoperability standard reduced the needs to re-enter geometric data, thermal values, and material properties (Kam et al 2002). Efficient model and data sharing during the schematic design phase allowed the building-systems consultants to simulate various cooling and heating requirements based on the architectural product model. These thermal values were directly imported to the Heating, Ventilating, and Air-Conditioning (HVAC) design application, where the mechanical consultants directed the distribution path. The software then automatically sized and balanced the mechanical components. The object-oriented HVAC application then exported the geometry of ducts and air handling units for the architects to incorporate into the architectural model and generated a bill of materials for the general contractor.

Synthesizing the readily available bill of materials from the mechanical consultants and the three-dimensional geometry from the architects, the general contractor for HUT-600 utilized an automated cost estimating and value engineering system, again object-oriented and IFC-compliant, to match design components with the contractor's database for cost estimation, scheduling, and resource leveling.

Compared to a conventional approach, this relatively seamless data exchange and the related automated tools tremendously expedited design time and improved the quality of interdisciplinary collaboration (Kam et al 2002). As a result, the project team quickly generated three design and two building system alternatives and was ready to take advantage of the time savings to perform further value-added work.

4.0 LIFE-CYCLE ANALYSES

The PM4D Approaches and PM4D Processes called for an enhancement of the facility performance over its total life span. Hence, HUT-600's project team conducted a series of life-cycle analysis to evaluate the thermal performance, cost implications, and environmental impacts of project alternatives.

In the following subsections, we explain how the project team evaluated two air-conditioning system alternatives—mixed cooling versus displaced cooling systems. In mixed cooling, the system supplies high velocity cold air from the ceiling. It is simpler in design and cheaper in cost when compared to a displaced cooling system, which slowly cools the space from the floor and displaces the warm air up to the exhaust in the ceiling.

[6] URL: http://www.graphisoft.com
[7] URL: http://www.progman.fi/english/e_index.htm

4.1 Thermal Performance

The mechanical engineers utilized Olof Granlund Oy's[8] RIUSKA[9] for thermal simulations and AEA Technology's CFX[10] to conduct computational fluid dynamics (CFD) analyses. Since the auditorium space was a critical room with heat emission from 600 users and more than 200 light fixtures, RIUSKA's predictions and CFX's analyses enabled the engineers to quantitatively compare the profiles of temperature and air velocity stratification between the mixed and the displaced cooling schemes.

RIUSKA accounted for the dynamic behavior of thermal masses in response to the changing exterior temperatures through an hourly increment over a 12-month period. Thus, engineers could combine different spaces and building systems to test various insulation and construction assembly options. Once the indoor air temperature target was specified, the program took a few minutes to analyze the thermal loads from the occupancy, the occupants' schedule, the equipment, and the exterior temperature conditions against the different insulation schemes, window transmittance, and louver systems.

Taking RIUSKA's analysis results as its target range and boundary conditions, CFX took about 10 hours to iteratively solve for the finite numeric values of air temperature and supply air velocity across the sectional profiles of the auditorium space. The CFX results illustrated that in spite of a supply of lower air temperature of 17 degree Celsius, the mixed cooling system was not as efficient as the displaced cooling system (required supply air temperature of 19 degree Celsius) in the occupants' zone—the area that mattered most. Hence, the engineers learned from numerical values and vivid graphical profiles (see Figure 1) that the mixed system had to supply cooler air at higher velocity in order to balance the warmer air around the lighting fixtures in the ceiling level.

4.2 Cost Analysis

The HUT-600 project consultants employed Olof Granlund Oy's BSLCC software to project the operation and maintenance costs of project alternatives throughout the facility's expected life-span. The consultants and the construction managers shared their respective knowledge from past projects, facility management data, and the manufacturers' catalogues to estimate energy consumption costs, maintenance costs, and immediate investment costs (see Figure 2), which all together provided reliable quantitative decision supports for selecting mechanical system, choosing electrical lighting and maintenance methods, and qualifying bid packages from air handling unit manufacturers.

In the mixed versus displaced cooling analysis, HUT-600 consultants assumed the systems had a 50-year service life span. For both alternatives, BSLCC read the automatically generated bill of materials from the object-oriented applications, estimated the cost of initial investment, and projected operation cost based on energy consumption and system efficiency. With the analysis tool, the consultants also accounted for the maintenance cost, replacement cost, financing cost, as well as inflation cost. The in-depth comparison results informed the decision makers that the equivalent annual cost of the displaced air-conditioning system was 6% higher than that of the mixed air-conditioning system.

4.3 Environmental Impact Analysis

With Olof Granlund Oy's BSLCA software, the building system consultants conducted environmental impact assessments to evaluate the environmental influences of the building materials and the estimated energy consumed by the facility. In particular, the consultants extracted the material properties and quantity information from the product model of the alternative designs. They deduced the level of environmental impacts to air and water and

[8] URL: http://www.granlund.fi/English/runko.htm
[9] URL: http://www.eren.doe.gov/buildings/tools_directory/software/riuska.htm
[10] URL: http://www.software.aeat.com/cfx

subsequently, they quantified the amount of pollution emission, global warming, acidification, etc. in support for material and system comparison (see Figure 3).

Iteratively, the designers, consultants, and construction managers evaluated and counter-proposed materials, structural systems, and building systems among themselves to balance aesthetics, performance, cost, and environmental impacts during the project design phase.

5.0 IMPLICATIONS

Kam et al (2002) reported that in HUT-600 the use of object-oriented product models and interoperability standard resulted in about 50% time savings in design documentation, as a result of object-oriented library, parametric properties, knowledge reuse, and data sharing. PM4D Approaches and PM4D Processes expedited the traditional schematic design services. As a result, the project team shifted their attention from performing routine jobs to conducting life-cycle analyses. Such analyses added project values while reducing the risks of cost overrun or dissatisfaction of long-term performance.

Had the HUT-600 project team not conducted life-cycle analyses on mixed versus displaced cooling systems, risk-adverse decision-makers would probably select the mixed system rather than the displaced cooling system. Since the latter system required an expensive under-floor distribution system, the lack of its performance evidences and operating cost information projection would have posted skepticisms to decision makers. While in HUT-600, not only did PM4D Approaches and analysis tools ensure the early availability of life-cycle analyses, they also improved the quality of the decision factors. For instance, the dynamic thermal behaviors that RIUSKA analyzed for the thermal performance are usually approximated or omitted in conventional design approaches.

HUT-600 consultants noted, and conferred by the project construction managers, that in the total spending on a capital facility, 80% of the total cost is spent on the operation and maintenance, whereas the remaining 20% goes to planning, design, and construction. Hence, it is crucial to capitalize on early project opportunities to optimize the facility design for long-term performance.

By March 2001, only three months into the design phase, the abovementioned analytical results were available to the owner and the project team members. Subsequently, the owners evaluated various project alternatives (e.g., architectural features, mixed versus displaced cooling systems, etc.) based on their functional performance, projected operating costs, maintenance costs, and environmental impacts and chose the most efficient designs and systems that best met their long term strategic goals. In the selection of air-conditioning system, the owners were confident to invest in the more energy efficient, environment-friendly, slightly more expensive, and better performance system—the displaced cooling system.

Furthermore, the HUT-600 project team also explored various visualization tools, such as virtual reality-EVE[11], 4D CAD[12], virtual animations, etc., to foster early and effective communications among the end-users, owners, and project team.

While all the HUT-600 project stakeholders enjoyed the benefits that we documented in this paper, there were still barriers and wish-list items that could have further improved the design services and better expedited the data sharing process. Such wish-list items included simpler data conversion processes, stronger interoperability, and better supports for revision handling (Kam et al 2002).

6.0 CONCLUSION

Shared among the owners, project team, and research partner of HUT-600 was a committed belief in capitalizing on early project

[11] URL: http://www.tml.hut.fi/Research/HUTVE
[12] URL: http://www.commonpointinc.com

opportunities to make a lasting and positive effect on the facility over its total life-span. Leveraging upon the reduced design time and improved data exchanges, enabled by the PM4D Approaches and Processes, HUT-600's project team was able to complete a series of life-cycle studies within the original design schedule. Pertinent decision factors and project alternatives were available during the early schematic design phase when making a decision had a relatively high impact and low cost.

In the nutshell, modern technologies are capable of expediting conventional design practices and promoting life-cycle approaches. Project experiences from HUT-600 construction pilot demonstrated that owners are empowered to choose among comprehensive life-cycle alternatives and to align the long-term facility values with their strategic plans, whereas project team members could differentiate themselves from their competitors with higher efficiency, better quality, and more effective application of their expertise.

7.0 REFERENCES

Kam, C. and Fischer, M. (2001). "Experiences from HUT-600 Auditorium Project." *Presentation at VERA Seminar*, Helsinki, Finland, November 20.

Kam, C.; Fischer, M.; Kunz, J.; Hänninen, R.; Lehto, S.; and Laitinen, J. (2002). "PM4D Final Report" *Working Paper*, Center for Integrated Facility Engineering, Stanford University, Stanford, CA, in press.

Paulson, B. (1976). "Designing to Reduce Construction Costs." *Journal of the Construction Division*, ASCE, Vol. 102, No. CO4, Proc. Paper 12600, December, pp. 587-592.

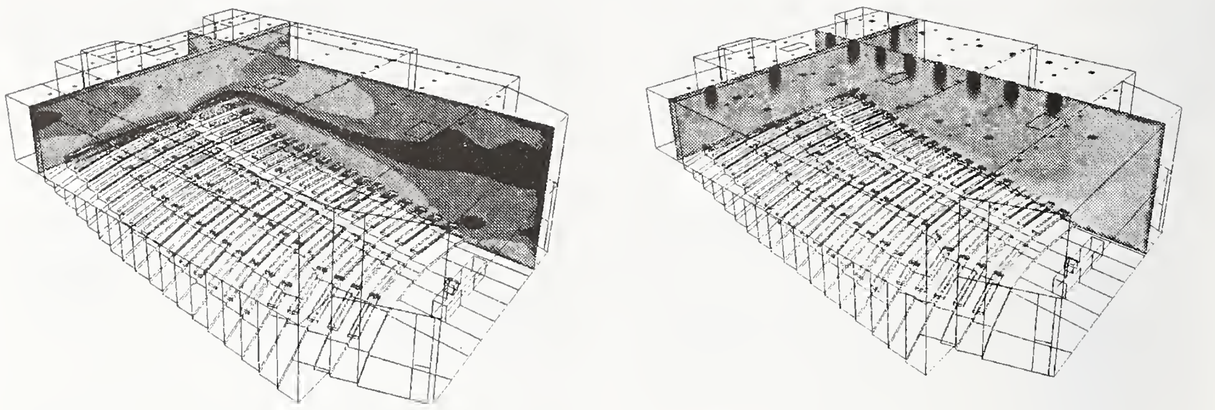


Figure 1. CFX provided CFD cross-sectional profiles of air velocity, which scales from 0.02m/s (blue) to 0.20m/s (red), in the displaced cooling scenario (Left) as well as the mixed cooling scenario (Right).

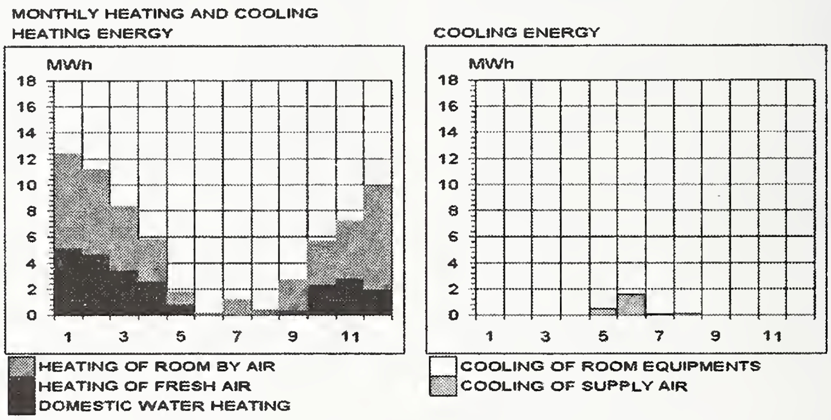


Figure 2. RIUSKA projected the annual heating and cooling energy consumption for HUT-600.

Level:	Property	Case 1:	Displacement ventilation (supply air from flo
Life Cycle Period,	50	Case 2:	Mixing ventilation (supply air from ceiling)

EMISSIONS TO AIR

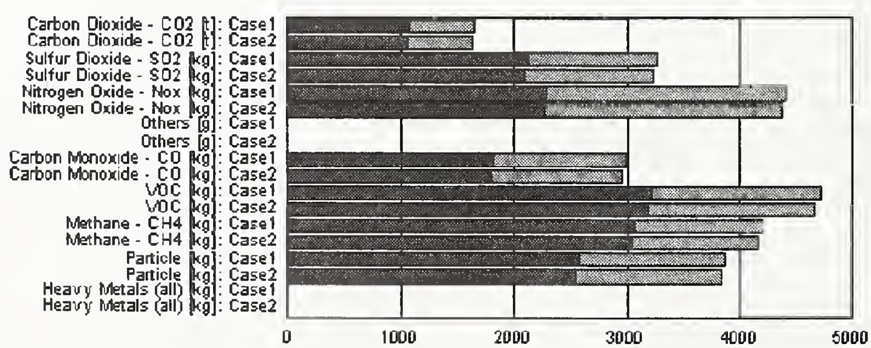


Figure 3. Charts from an environmental impact analysis showing the weight of emission (units in y-axis brackets) from energy consumption (blue) and the building materials (green) over a 50 year period in HUT-600.

SESSION 2

CONSTRUCTION PROCESS MODELING AND SIMULATION

Dynamic Change Management for Fast-tracking Construction Projects

by

Moonseo Park[1]

ABSTRACT: Uncertainties make construction dynamic and unstable, mostly by creating non value-adding change iterations among construction processes. Particularly, when a project is fast-tracked without proper planning, those change iterations can cause the disruption of the construction process. For this reason, to effectively handle fast-tracking change iterations involved in fast-tracking need to be identified, and the dynamic behavior of construction resulting from those change iterations must be dealt with in a systematic manner. As an effort to address some of these challenging issues in fast-tracking construction, this research paper identifies different change iteration cycles involved in fast-tracking construction and observes the characteristics and behavior patterns of change. All of research findings are incorporated into a cohesive system dynamics model and the model simulation confirms that managerial decisions on change or rework should be made based on the proper assessment of their tradeoff. In addition, a case study of highway and bridge construction projects shows the potential of how fast-tracking construction can benefit from dynamic change management in real world settings.

KEYWORDS: Change, Fast-tracking, Feedback, Rework, Simulation, System Dynamics

1. INTRODUCTION

Shortening time-to-market has been one of the most critical factors to the success of businesses in many industries. As a result, companies have sought a method that can ensure a faster product development, most commonly focusing on product cycle time reduction through concurrent development. The construction industry is not an exception. The increasing preference of project owners and managers to fast-track construction proves the popularity of concurrent development in construction. In addition, many success stories of fast-tracking have demonstrated that the popularity of this delivery method is warranted [Huovila et al., 1994; Williams, 1995]. However, concurrent construction also has greater potential to impact the project development process than the traditional more serial method [Pena-Mora and Park, 2001].

In the literature, these potential problems are mainly attributed to the increased level of uncertainty and research efforts on fast-tracking have focused on uncertainty reduction. However, in dealing with uncertainties, most of the previous researches have not explicitly addressed how they impede construction processes, nor identified the different patterns of their impact on the project performance.

Closer observations of the design and construction process indicate that uncertainties make the construction dynamic and unstable, mostly by creating non value-adding change iterations among construction processes. Particularly, when a project is fast-tracked without proper planning, those change iterations can cause the disruption of the construction process. In addition, people's preference of change to rework can reinforce the change impact. Since construction has a physical manifestation, construction rework is normally perceived to have a bigger impact than change. As a result, construction managers tend to avoid rework on problematic tasks by changing the scope of work, in particular under time constraints. However,

[1] Assistant Professor, Department of Building, School of Design & Environment, National University of Singapore, Email:bdgmp@nus.edu.sg

construction sequence by triggering subsequent changes on other tasks, which often contributes to unanticipated schedule delays and cost overruns in fast-tracking construction. For these reasons, to effectively handle fast-tracking change iterations involved in fast-tracking need to be identified, and the dynamic behavior of construction resulting from those change iterations must be dealt with in a systematic manner.

2. CONSTRUCTION CHANGES

Non-value adding iterations in construction are mainly associated with construction changes. Accordingly, reducing wasteful construction iterations requires effective change management, which should start with the understanding of different characteristics and behavior types of construction changes.

2.1 Types of Changes

Normally, construction changes refer to work state, processes, or methods that deviate from the original construction plan or specification. They usually result from work quality, work conditions or scope changes. Meanwhile, changes that have been already made (denoted as Changes as Result in Figure 1) can be the source of subsequent changes in other tasks (denoted as Changes as Source in Figure 1). For example, changes in the design work that have been made by mistake can cause subsequent changes in construction. In this case, the design changes are a result to the designer, while they can be a need for changes to the construction crew. In addition, change can be also seen as an action of making a change (denoted as Change as Behavior in Figure 1), which is further categorized into unintended change and managerial change. Unintended changes occur without the intervention of managerial actions. The arrows labeled *E*, *F*, and *G* in Figure 1 illustrate the unintended change process. Meanwhile, managerial changes are made by managerial decisions during quality management or project monitoring and control. As illustrated in Figure 1, once changes occur during construction (*A* and *B*), changes result in either subsequent changes (*C*) or rework (*D*), depending on managerial decisions.

2.2 Differentiating from Rework

Both change and rework are done in the form of either 'adding', 'deleting' or 'replacement (deleting and adding)'. However, given the same problem, they have different behavior patterns, since change and rework have different characteristics, as summarized in Table 1. For example, in Case I on Figure 2, given the problem (a hump on the concrete surface), rework would be done by deleting the problem, while change would be done by adding some more concrete. In addition, in Case IV where floor tiling has been finished with less than the required height, although both change and rework have the same behavior pattern (replacement) in solving the problem, the object would be the problem area in rework, while the previous work would be the object in change.

2.3 Tradeoffs

In construction, the change option is more general. Since construction has a physical manifestation, construction rework is usually accompanied with the demolition of what have been already built, which normally has a bigger direct impact on the construction performance than the change option. By adopting the change option, it is possible to avoid rework on problematic tasks that may require more resources. However, as previously discussed, changed tasks can also become a change source that can cause other subsequent changes, which might have more impact on the construction performance than the rework option in certain conditions. For example, the increased concrete height in Case I and Case III on Figure 2 may trigger subsequent changes in succeeding tasks, i.e., reducing the size of ventilation ducts. In addition, in Case V on Figure 2 where some of piles have not been correctly positioned, it may be possible to proceed with the superstructure without correcting the position of the piles by changing the position of columns. However, this change option may necessitate unplanned cantilever construction in order to keep the original floor layout, which needs to be evaluated as compared to re-driving the piles. Consequently, a decision on the change option needs to be carefully made based on a good understanding of how changes evolve to non-value adding iterations, which can create unanticipated and indirect side effects of the decision.

3. DYNAMIC PROJECT MODEL

The dynamic project model to be presented has been developed taking into consideration effective change management and operation level construction policy making. To develop the model feedback processes involved in fast-tracking construction were identified focusing on how they can trigger non-value adding iterations in the form of construction changes. Having identified feedback processes, the generic construction process, which constitutes the skeleton of the project model, was modeled.

3.1 Feedback Processes in Construction

Normally, construction involves feedback processes represented in the causal loop diagram on Figure 3-a, 3-b, 3-c, and 3-d. When tasks and resources are available, first, the upstream work, based on which the available tasks will be carried out, is reviewed before commissioning resources for the tasks. During the review process, problems made in the upstream work can be discovered. Once they are found, depending on managerial decisions, workers may request the upstream worker to correct the problematic work. More upstream hidden changes can cause more requests for the upstream work reprocess, which results in more pending tasks (*A*) and schedule delays (*B*) in the downstream work. Otherwise, workers construct tasks not having problems in the associated upstream tasks, with given resources. Once tasks are completed, the construction performance on the tasks is periodically monitored or inspected to see whether or not the target quality is met and the intended functions are achieved. Through this quality management process, a decision on whether releasing the completed tasks or not can be made.

Unintended changes resulting from low work quality, bad work conditions or frequent scope changes can cause managerial changes (*C*), rework (*D*), or hidden changes (*E*), depending on managers' willingness to adopt the change option and quality management thoroughness. The more construction is delayed the more often the change option tends to be adopted (*F*), in order to avoid rework, which is normally perceived to have a bigger impact on the schedule performance. However, such managerial efforts can create unplanned and/or indirect side effects. As a result of feedbacks involved in the processes (*F, G, H, I, J*), managerial changes can trigger further delays as well as rework. As diagramed in Figure 3-a, managerial changes trigger reprocess iterations of tasks that have been already released (refer to the definition of managerial changes in Table 1), while rework delays the construction progress by creating reprocess iterations of tasks that have not been released.

In addition, delays also may make quality management efforts less thorough (*K*), which results in more hidden changes (*L*). During the downstream review process, hidden changes released from the upstream work can be discovered. Once they are found, depending on managerial decisions, downstream workers request the upstream worker to correct the hidden changes. As a result, more hidden changes can cause more correction requests from the downstream (*M*), which also can delay the construction progress as a result of subsequent feedback processes (*N, I, J*) diagramed in Figure 3-b.

Furthermore, increased willingness to adopt managerial changes also can increase subsequent changes in the downstream work (*O*), which delays the downstream work process. Consequently, reprocess requests from the downstream work are also delayed (*R*), which again impacts the schedule performance of the activity that has originated changes (*N, I, J*). Meanwhile, lowered quality management thoroughness creates more hidden changes (*L*). Increased hidden changes can deteriorate the work quality of the downstream work, which creates more reprocess iterations of the downstream tasks. This also impacts the upstream schedule performance through (*R, N, I, J*). All of these feedback processes can impact the construction performance, combined with resource availability, construction policies, and people' reactions to work conditions and policies.

3.2 Model Description

Based on feedback processes and relationships among construction variables in the causal loop diagrams on Figure 3, the quantitative representation of the generic construction process has been modeled using system dynamics modeling techniques. In addition, other supporting model structures for resources, scopes, and quality have been also developed. Detailed model descriptions are found in Park (2001).

4. CASE STUDIES

The developed dynamic project model is being applied to the construction of 27 bridges in order to help effectively manage changes and prepare a robust construction plan. The construction is a part of a $400 million Design/Build/Operate/Transfer project for roadway improvements along State Route 3 from its intersection with State Route 128 in Burlington, MA north to its terminus at the New Hampshire border. The development process is expected to span 42 months with the project completion achieved in February, 2004. The project scope includes widening the 21-mile of the state roadway and the existing 15 underpass bridges, and renovating 12 overpass bridges. This paper presents a case study of the Treble Cove Road Bridge Construction, one of bridge renovation projects, demonstrating how the case project has suffered from changes and providing construction policies to minimize change impact on the project performance including labor policies and schedule buffering.

4.1 Simulating the Actual Performance

The simulated actual duration is 559 working days. This is 168 days longer than the CPM-based duration of the base case, which is 391 working days. The difference in the completion time is mainly caused by a lot of non-value adding iterations among design and construction activities. Actually, the construction team is working to address such issues that the design development of the Treble Cove Road Bridge project was already shown significant delay and construction has not been yet started. Some of these issues are due to the fact that this project was awarded to the contractor before the detailed scope of the project has been established. As a result, changes on the design work were frequently requested from the owner side during sketch plan, final plan, and shop drawing submittal, which resulted in a lot of design iterations. In addition, this case project was the first design/build contract for the members of development team in the owner side, expected level of coordination among the owner, designer and constructor has not been met to date and design iterations encountered were difficult to handle. Based on interviews with the design and construction team, these challenges in the design development were represented as 'Highly Unreliable' in the project model and the simulated actual durations for those activities show how much non-value adding iterations caused by changes can affect the project progress in a quantitative manner.

4.2 Policy Implications

In order to examine the effectiveness of different construction policies, simulations were done adapting the actual case with different scenarios for managerial decisions on change or rework, labor control, buffering, and some important time variables. As a result of the simulations, the following policy implications were obtained (refer to Figure 4 to see the model simulation).

First, a higher managerial change ratio tended to reduce costs but lengthen the project duration. However, it is hard to generalize this result, since the tradeoff of change and rework is highly dependent on construction system conditions at the time when a decision is made. This implies that effective change management requires an operational level approach rather than a long term policy, and it should be accompanied with well preparation of relatively long-term policies such as labor control policies, schedule buffering and delivery methods.

In connection with labor policies, flexible labor control was found to be effective for the case project in terms of schedule and cost reduction. In contrast, overtime contributed to facilitating the project schedule to some extent but its effectiveness is questioned, once increased project costs are considered. Overtime applied for the case project lowered productivity and increased change rate, as workers' fatigue was accumulated. In fact, the effectiveness of labor control policies can vary depending on the nature of a project. However, many success stories of concurrent construction projects like our case project confirm the above policy implications, demonstrating that having flexibility in labor control contributes to reducing the project duration and costs by assigning workforce in a timely manner.

In addition, the case project has been simulated with various buffering scenarios; not having buffer, having uniform buffer, and having buffer based on activities' characteristics. The simulation results showed that applied buffers contributes to reducing the upstream change impact and non-value adding iterations. As a result, the resource idle time and waste were reduced, which made it possible to more effectively utilize given workforce. In particular, buffering based on activities' characteristics turned out to have most effectively enhanced the schedule and cost performance.

Lastly, the simulations done with different time variable scenarios demonstrate that shortening a required time for labor hiring and RFI reply contributes to enhancing the project schedule and cost performance. In particular, RFI reply time greatly affected the project performance. Shortening RFI reply time by half could facilitate the project progress by 12% and reduce the project costs by 10%. In contrast, when RFI reply time was doubled, duration and costs were increased by 29% and 24% respectively. These simulation results imply that for this case project, coordination among the project functions is crucial to the success of the project. Consequently, the decision-making process in design and construction should be shortened and information flow among project functions should be streamlined to assist in reducing the decision-making time.

In conclusion, although the obtained simulation results can vary depending on project settings, they well demonstrate how the dynamic project model can contribute to enhancing the project performance in a real world setting by providing effective change management plans and policy guidelines. Additionally, the simulation results also imply that model-based construction policies can be more effective, when combined with other managerial efforts such as reducing a process time and increasing the level of coordination among project functions.

5. CONCLUSIONS

Construction involves a lot of non value-adding change iterations due to its structural problems, in particular when construction is performed concurrently. This has necessitated the development of a tool that can effectively manage construction changes. This paper addressed the challenging issue by introducing the concept of dynamic change management to construction planning and management. Although the research results discussed thus far need to be further refined and developed, they demonstrated that the dynamic change management approach and the developed project model would help prepare a more robust construction plan against uncertainties and provide policy guidelines, by taking into consideration the context in which a construction project is being developed.

6. REFERENCES

Fazio, P., Moselhi, O., Theberge, P. and Revay, S. (1988), *Design Impact of Construction Fast-Track,* Construction Management and Economics, Vol. 6, No. 2, pp 195-208

Ford, D and Sterman, J. (1997), *Dynamic Modeling of Product Development Processes,* Sloan School of Management, Working Paper 3943-97, MIT, Cambridge, MA

Huovila, P., Koskela, L., and Lautanala, M. (1994), *Fast or Concurrent: The Art of Getting Construction Improved*, Proceedings, pp 143-158, The second workshop on lean construction, Santiago

Kwak, S. (1995), "*Policy Analysis of Hanford Tank Farm Operations with System Dynamics Approach"*, Doctoral Thesis, Dept. of Nuclear Engineering, MIT, Cambridge, MA

Lyneis, J. (1999), "Dynamics of Project Performance", Course Material, Dept. of Civil and Envr. Eng. at MIT

Park, M. (2001), "Dynamic Planning and Control Methodology for Large-Scale Concurrent Construction Projects", Doctoral Thesis, Dept. of Civil Engineering, MIT, Cambridge, MA

Pena-Mora, F and Park, M. (2001), "Robust Control of Cost Impact on Fast-tracking Building Construction Projects', Journal of Construction Engineering and Management, ASCE

Tighe, J (1991), Benefits of Fast Tracking are a Myth, International Journal of Project Management, Vol. 9, No. 1, pp 49-51

Williams, G.V. (1995), Fast-Track Pros and Cons: Considerations for Industrial Projects, Journal of Management in Engineering, Vol. 11 No 5, pp 24-32, Sep/Oct. 1995

Figure 1: Changes as Iteration Trigger

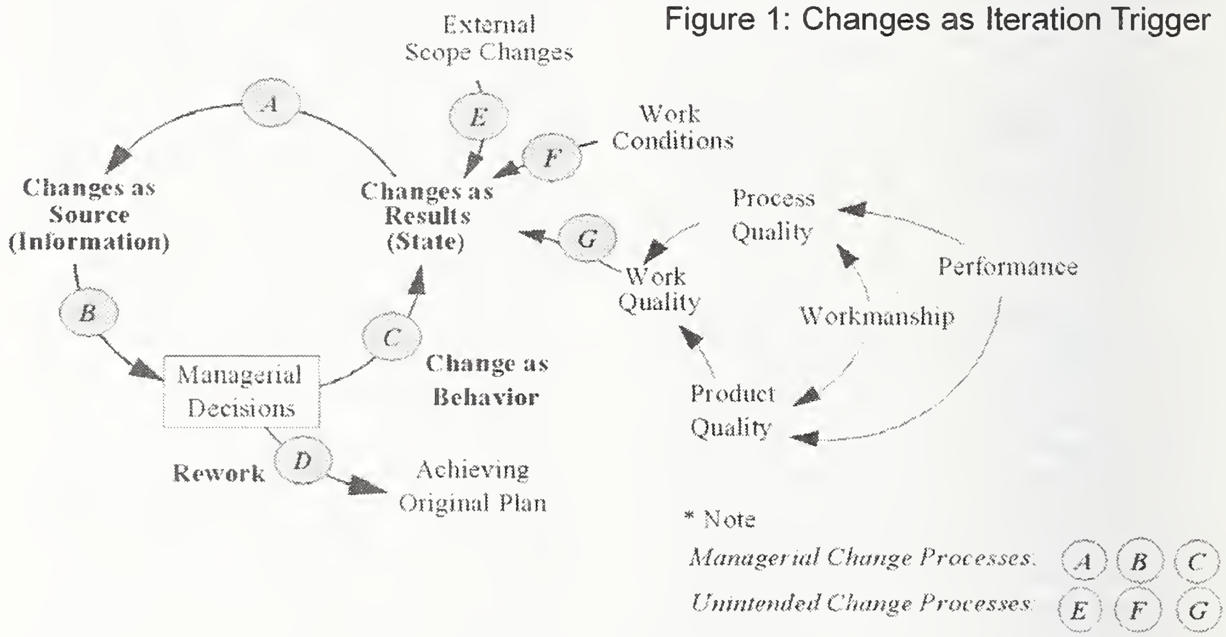


Table 1: Characteristics of Change and Rework

Des.	Purposes	Object	Scope of Work	Triggering Another Changes
Managerial Changes	Minimizing the impact of changes that have already occurred by adopting a different method or process than in the original plan and specification.	Preceding or succeeding tasks	Vary, depending on sensitivity	May trigger
Rework	Achieving what are originally intended in the plan and specification.	The problematic tasks	Same as the scope of problematic tasks	None

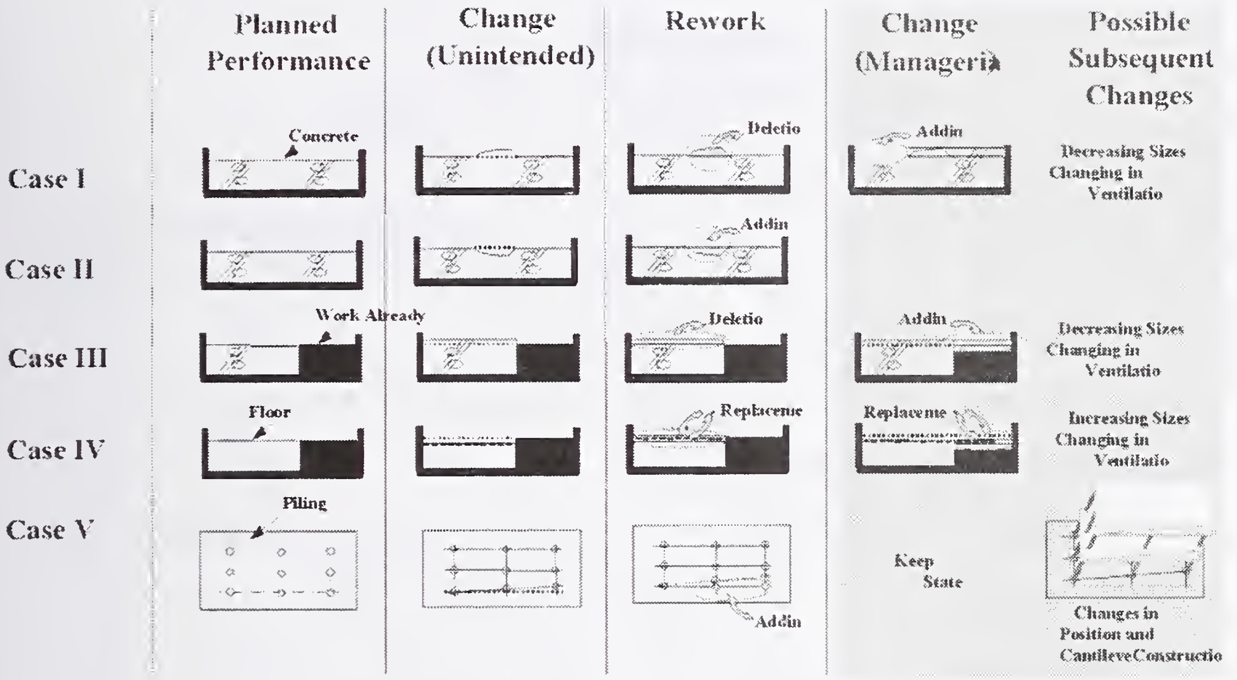


Figure 2: Behaviors of Change and Rework

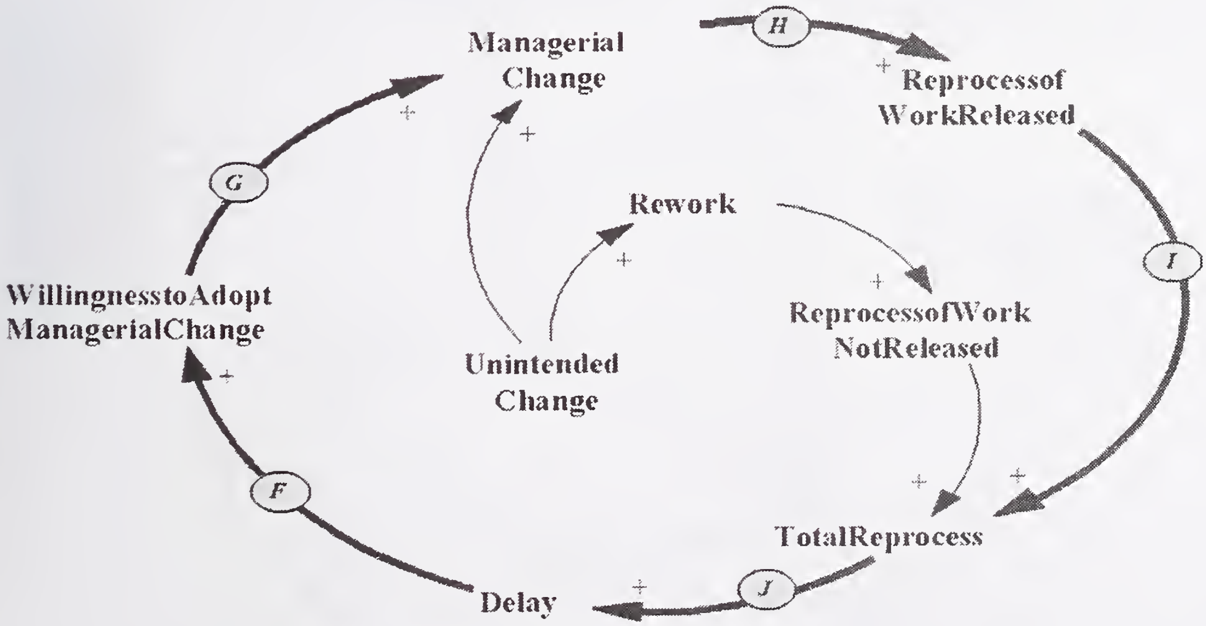


Figure 3a: Change Option Loop

Figure 3b: Quality Management Thoroughness Loop

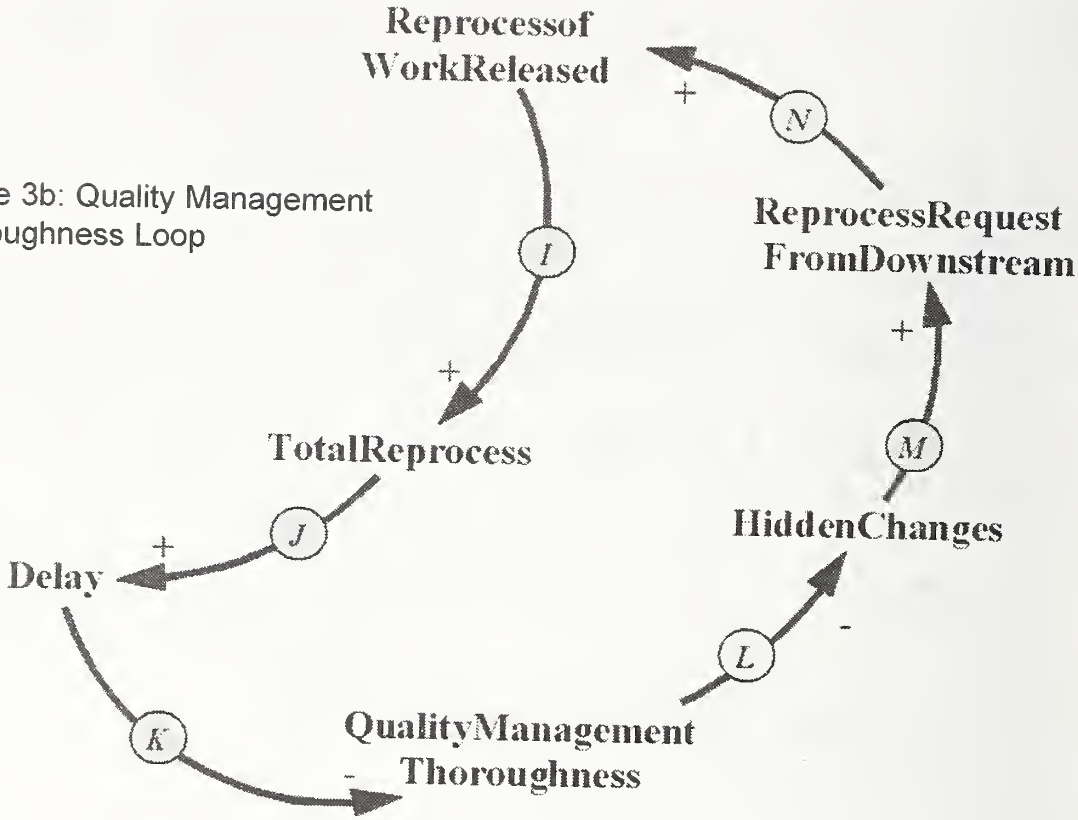
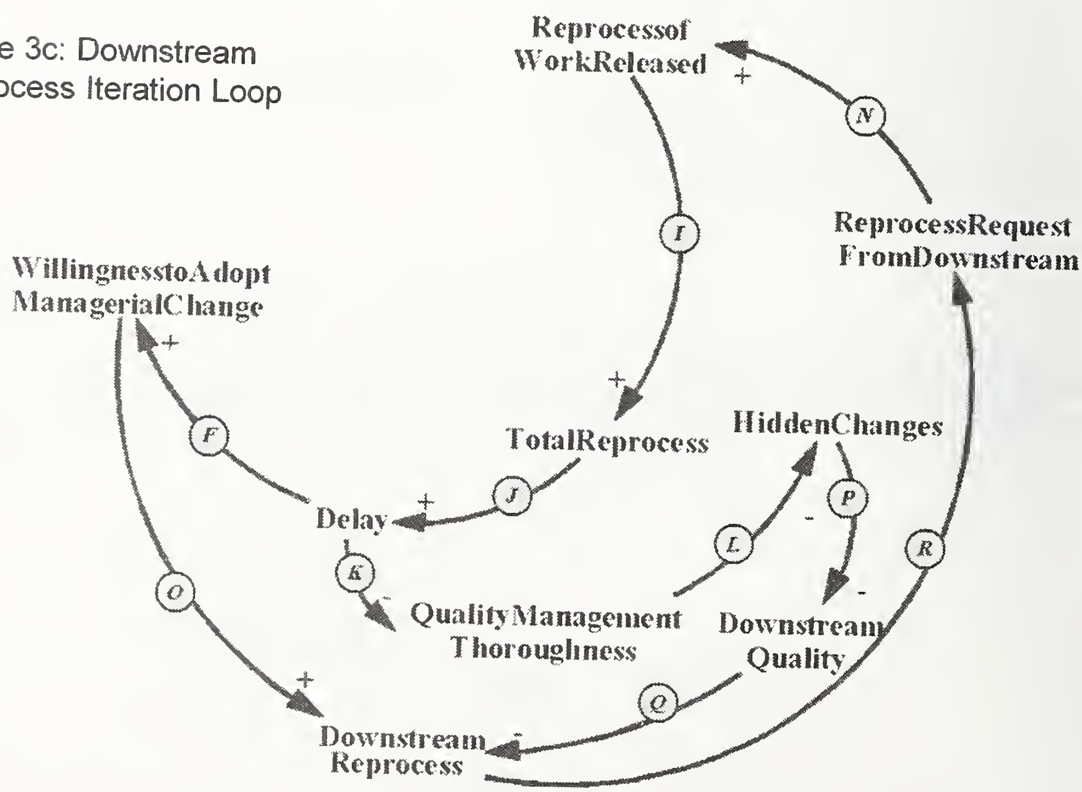


Figure 3c: Downstream Reprocess Iteration Loop



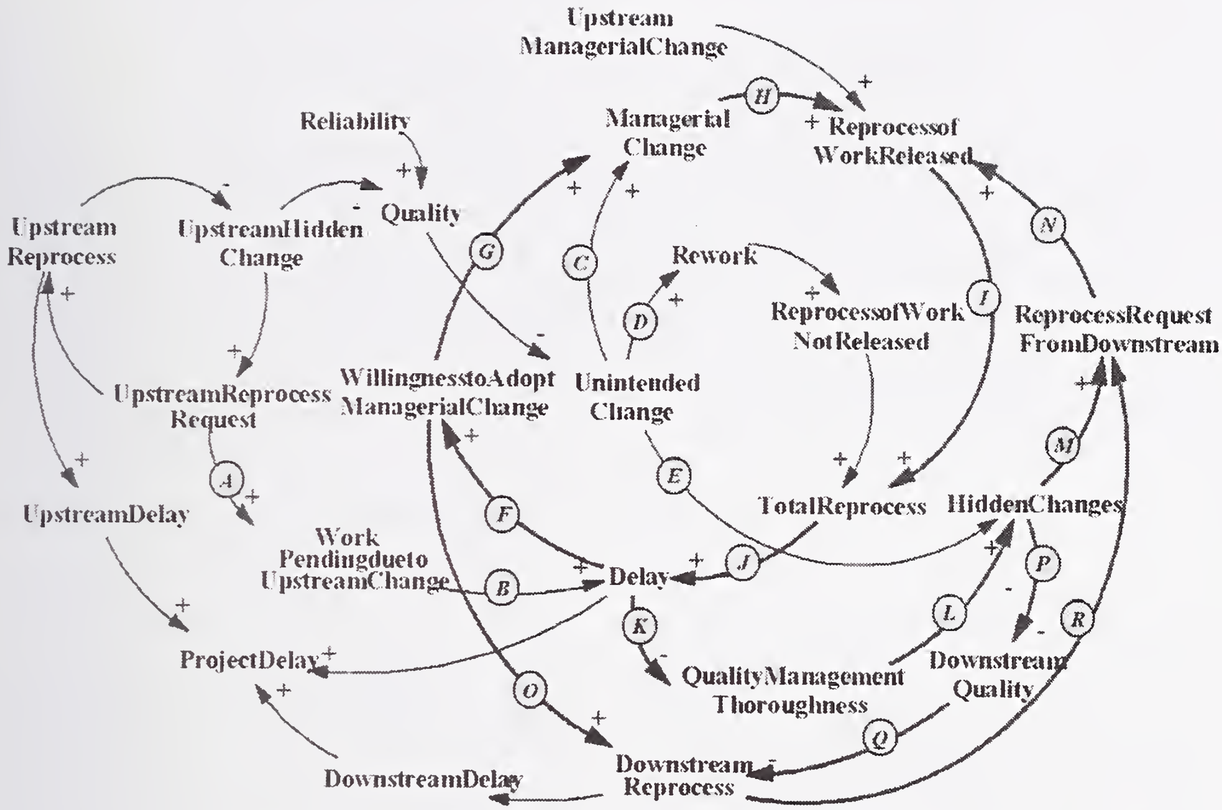


Figure 3d: Feedback Processes in Construction Activities

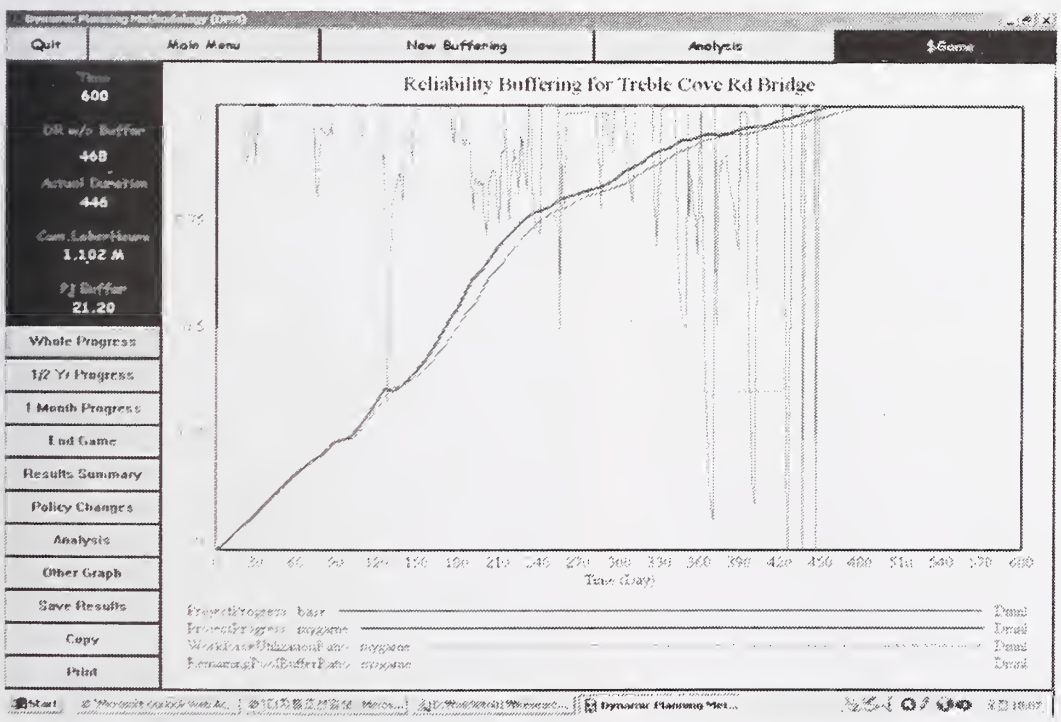


Figure 4: The Case Project Model Simulation

AUTMOD3: A PLANNING TOOL FOR MODULAR BUILDING SYSTEM

by

V.M. Padron, O. Cardenas, R. Diez, M. Abderrahim, C. Balaguer [1]

ABSTRACT: High quality modular construction is one of the solutions for the fast growing need for affordable high quality housing in Europe, which can not be solved by conventional building technology. The software environment for automatic modular construction AUTMOD3 has been developed, by University Carlos III de Madrid, in the frame of the European Union project FutureHome. The environment integrated in the Computer Integrated Construction (CIC) concept is composed by several tools: design, planning and simulation, which are linked and able to interchange their data. The tools have been integrated in a well known CAD system and the their work has been tested on an automatic crane for the assembly of pre-fabricated modules on the construction site. This paper describes the AUTMOD3 planning tool.

KEYWORDS: Automatic modular construction, CAD/CAM Integration, CIC, Robotics

1. INTRODUCTION

FutureHome is an European Union funded research project, which is aimed to the development of high quality modular construction and its required Information Technology infrastructure for the european construction industry. High quality modular construction is one the solution for the fast growing need for affordable high quality housing in Europe, which can not be solved by conventional building technology. Modular construction states building process as the assembly of pre-fabricated modules i.e. initially a modular design of the construction project is obtained, then these modules are produced in a factory, and finally, they are assembled on the construction site. The use of pre-fabrication leads to the following advantages: a) structuring of the construction process allows to apply automation, robotics and computer integrated construction methods (computer-aided design, planning, control, supervision tools and automation systems), b) reduction of the health risks of workers and improving their working conditions, c) increasing building process predictability (building process is less affected by weather conditions, there is a greater control on materials use and the supply chain) and d) decreasing waste and increasing productivity. The participants of the project are steel and building companies, research centres and universities from Finland, Sweden, Germany, The Netherlands, United Kingdom and Spain. In the frame of the FutureHome project Carlos III University of Madrid has developed an automatic modular construction software environment, AUTMOD3, realised in the Computer Integrated Construction concept [1,13]. The system consists of a design tool [5], a planning tool and a simulation tool. This paper presents the planner tool, which links the design tool with the simulation and the execution of the real assembly employing a robotized crane, it takes as input a finished design and obtains the sequence in which modules should be produced, transported and assembled on the site, as well as the robotized crane commands needed for the automatic assembly of the building. The second section of this paper presents the main concepts, the state-of-the-art and the solutions applied for the assembly sequence and the motion planning, third section shows the work of the planner and the results of its application on the robotized crane, finally conclusions are presented.

2. AUTMOD3 PLANNING TOOL

The production sequence and the on-site modular assembly planning are the objectives of the AUTMOD3 planning tool. This planner will define the sequence, in which pre-fabricated modules should be produced, transported and finally assembled on the construction site. Also, it will generate programs for the automated devices involved in the on-site assembly.

[1] Robotics Lab, Carlos III University of Madrid
C/ Butarque 15, 28911 Leganés, Madrid (Spain)
e-mail: vmpadron@ing.uc3m.es

Assembly planning is a very complex problem when it is viewed in the context of robots and other automated devices programming (Figure 1). Assembly planning defines the assembly operations and their sequence, the fixture design, the tool selection and the workspace layout, satisfying a set of technical and economical requirements. This process produces the assembly sequence, which is executed by an automated device. In order to do this, it is necessary to plan the grasp points selection and the grasping of each module, the path planning from the module initial position to the destination position and the execution of the fine motion to insert the module, making a suitable use of the device sensors during all these operations [14].

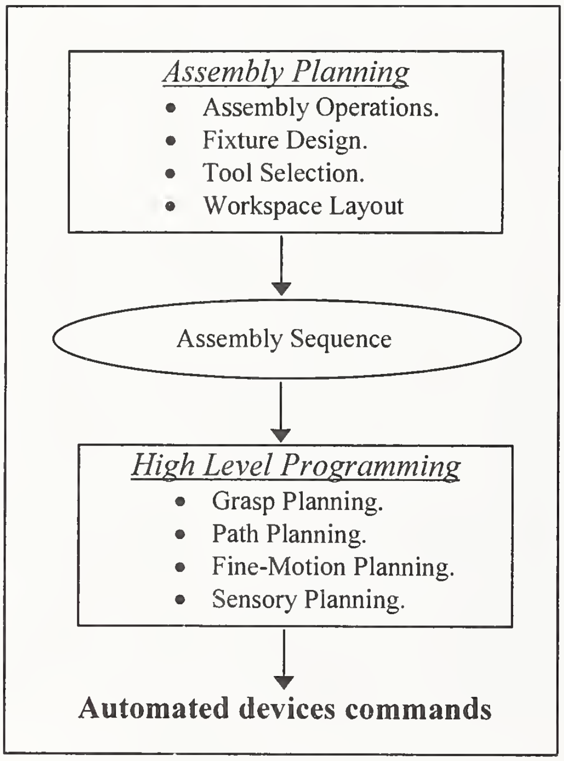
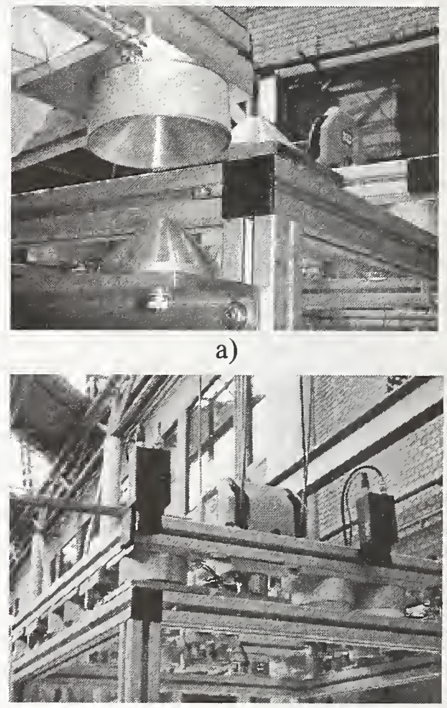


Figure 1. Assembly planning in the context of automated devices programming.

A key concept to obtain a successful automated assembly system is the integration of the planning process with the design and execution processes [2]. This concept has been applied in FutureHome. On one hand, specially developed assembly connectors has been added to the pre-fabricated modules to allow a bigger tolerance of robotized crane movements and to help the insertion of modules (Figure 2-a). On the other hand, the crane has been provided with a platform carrying electromagnets and an artificial vision system that allow to locate, grasp and insert modules in an automatic way (Figure 2-b). This approach permits the simplification of the high level programming allowing to decompose assembly planning in two simpler processes: assembly sequence planning and motion planning. Assembly sequence planning searches the order and spatial direction in which the insertion of each module should be done. Motion planning defines the spatial trajectories of the modules during the assembly and the automated devices commands needed for their execution.



a)

b)

Figure 2. a) Cones connector, b) platform artificial vision system.

2.1 Assembly sequence planning

Assembly sequence planning searches the order and spatial direction in which the insertion of each module should be done. This sequence should satisfy the following types of restrictions and criteria: a)geometrical restrictions, b)physical restrictions, and c)assembly process related restrictions and criteria. Geometrical restrictions are those related to the collision between parts in the assembly process. Physical restrictions are those related to the unwanted changes due to gravity, unwanted motion, etc. (these restrictions are also defined as the assembly stability). The assembly process (available fixtures, tools and workspace) can be specified to the assembly sequence planer by means of a set of restrictions and criteria. A criterion, unlike a restriction, is a measure for evaluation. This kind of restrictions and criteria allow the generation of different sequences and the selection of the most suitable one.

Assembly sequences search is a NP-complete computational problem [8], that is the time needed to determine the optimal assembly sequence grows exponentially with the number of parts. Due to this complexity most of the existing planners are interactive systems, which generate two-handed monotone assembly sequences in reverse, starting from the more highly constrained, full assembly state [15]. Some of the most known assembly planners are: *DFA 8.0 and DFA/Pro*, this is a very useful advisory system, widely used in the industry, and *Archimedes 4.0* an ACIS enabled automatic planner with an interactive user interface and a library of restrictions, that allows assembly planning in a very pragmatic way. The authors of the system have reported the assembly planning of a Hughes Aircraft air-to-air missile guidance section with 472 parts in 22 minutes [6].

In the FutureHome project, pre-fabricated modules are vertically assembled using a cones connector system (Figure 2-a). This system allows the auto-centring of the modules during the insertion, as well as a bigger tolerance in the precision of the crane movements. The assembly by means of cones connectors can be considered a stack assembly (i. e., one in which all the assembling motions occur in a single direction, usually up-and-down.). In a construction work, assembly is always performed on a structural basis. This leads to sort assembly sequence inserting first those modules of bigger "structurality" i.e. modules that support another modules. This sort according "structurality" is performed on levels, and inside the same level, it is performed on co-ordinates. Therefore, if these attributes: "structurality", level and co-ordinates of the insertion point, are assigned to each module, then it is only needed to sort the modules according to these attributes to obtain the assembly sequence. In addition, a special attribute, grouping, is provided to take into account custom planning criteria. This attribute has the biggest priority allowing the division of a given work in smaller tasks. The sorting time for a building containing 70 modules using this method is instantaneous. The tool is also provided with a manual sorting feature to face unexpected events during the execution process.

2.2 Motion planning

Motion planning defines the spatial trajectories of the modules during the assembly and the automated device or robot commands needed for their execution. Motion planning is defined as follows: given the initial and the final configuration of the module, find a joining path that avoids collision with the obstacles. There are many techniques and algorithms for motion planning, but they are generally classified in three groups: roadmaps, cells decomposition and potential fields [10].

Though motion planning has been solved exactly [3], the obtained algorithms are slow and complex. The complexity of the motion planning problem grows exponentially with the number of degree of freedom (DOF) of the robot and in a polynomial way with the number of obstacles. Nevertheless, there are many problems in practice that are not really complex and that can be solved exactly or in a approximated way. For robots with four or less DOF of freedom it is possible to solve motion planning problem, using roadmaps or approximated cells techniques in C-space, within a few seconds or a few minutes, depending upon problem complexity [7]. Motion planning for robots with more than four DOF is an active research area. Some of the outstanding motion planners in this area are: the Probabilistic Roadmap Planner (PRM) – this planner generates a roadmap based on random, but properly selected, collision-free configuration nodes, and uses this roadmap to find paths [9], SANDROS – a hierarchical, multi-resolution dynamic-graph search planner [4] and AMROSE – a multi-agents, artificial potential based system applied to off-line high level programming of welding applications [12].

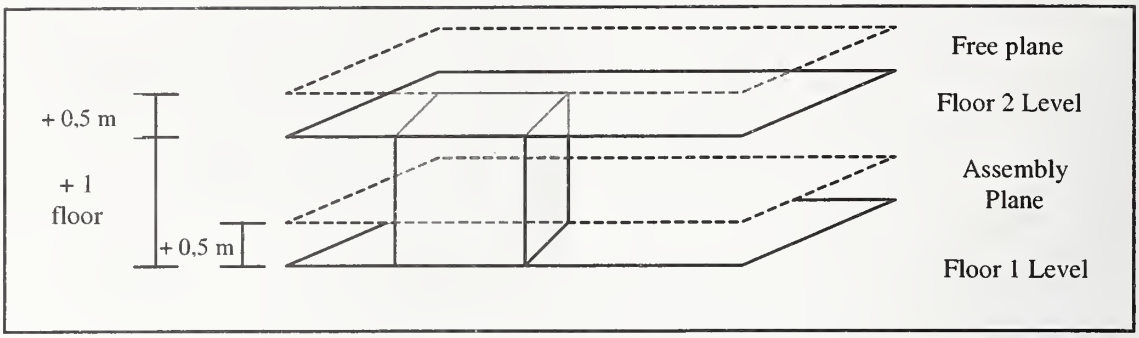


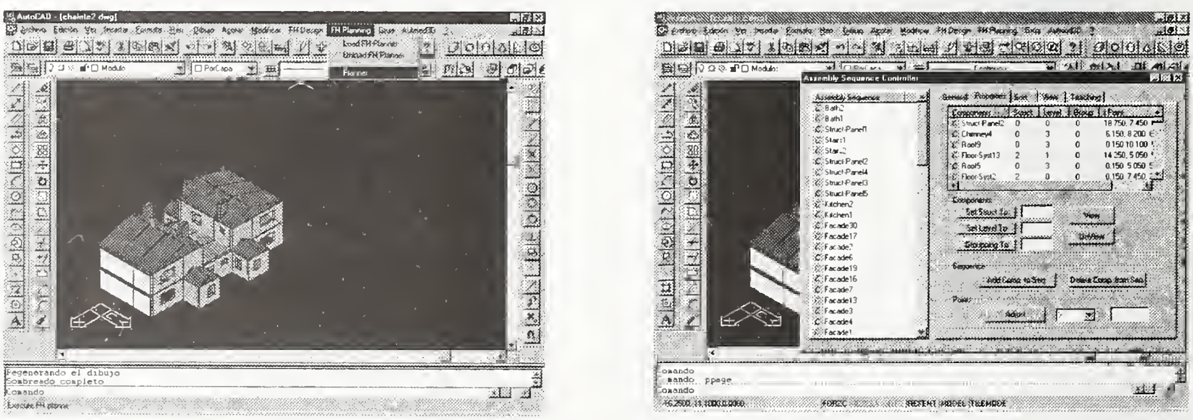
Figure 3. Working planes of the AUTMOD3 motion planning tool.

In order to reduce the motion planning complexity, AUTMOD3 motion planner basically works on two planes. These planes are called assembly plane and free plane (Figure 3). Assembly plane is defined at 0,5 m of height with respect to the current assembled level. Free plane is defined at 0,5 m of height with respect to the next (higher) level than the currently being assembled. Initially, the planner looks for a path free of collisions in the assembly plane using a variant of the tangent graph algorithm [4]. If this path does not exist, its distance is very large or the execution time of the algorithm is large, the planner rises the module to the free plane, carries it directly over the insertion point, lowers the module and proceeds to the insertion. In addition, the tool is provided with a manual planning function allowing the modification of the obtained trajectories or the generation of other types of paths. This method trades-off optimality versus complexity producing a simple algorithm.

Grasping operations are automatically performed by means of the electromagnets and the artificial vision system installed in the crane platform, while insertion operations are also automatically realised using the cones connectors and the crane platform artificial vision system.

3. APPLICATION TO THE FUTUREHOME ROBOTIZED CRANE

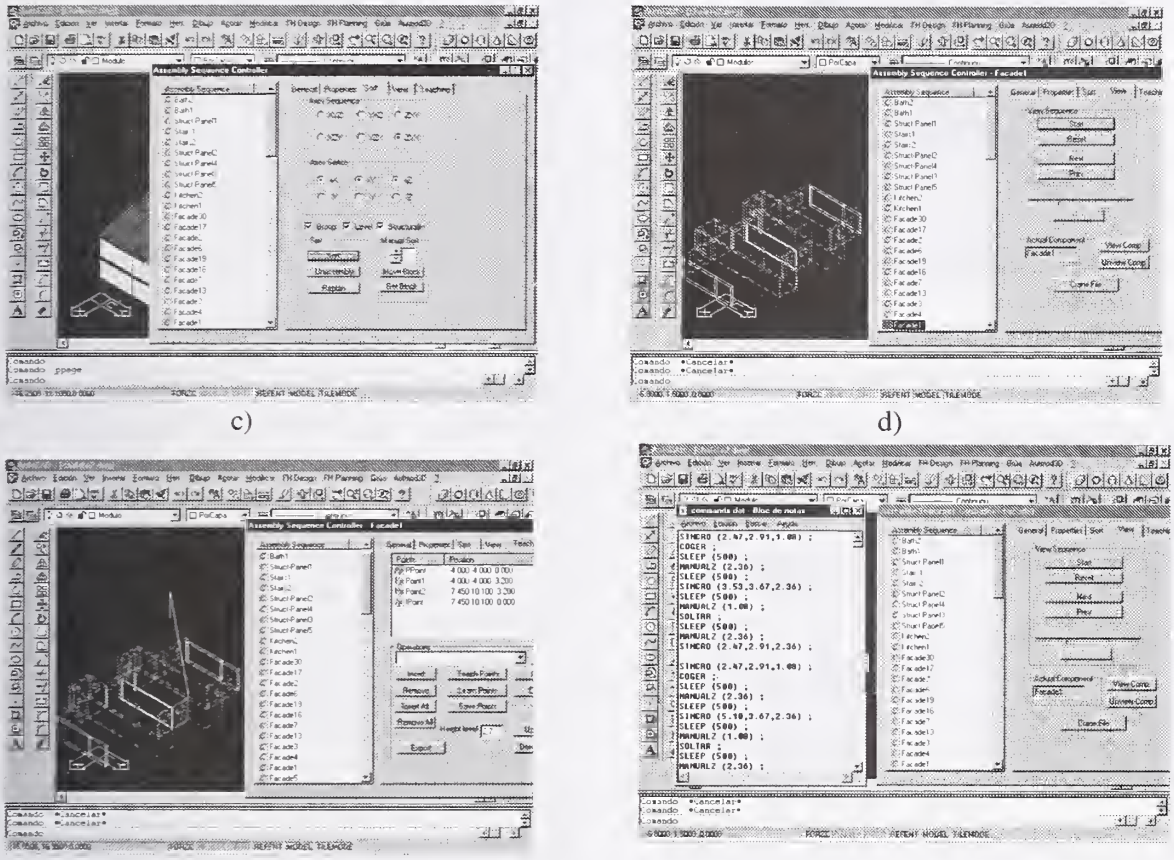
The AUTMOD3 assembly planning tool is integrated in the AutoCAD environment. It is accessible from the menu bar and communicates with the design tool through the AutoCAD database.



a)

b)

Figure 4. a) Calling the AUTMOD3 assembly planner, b) visualising modules assembly properties.



c) d) e) f)

Figure 4 (cont.). c) Sorting modules according to their properties, d) visualising the assembly, e) automatic motion planning: a trajectory for the assembly of the Facade 1 module, f) visualising robotized crane commands.

The planning tool is presented as a single dialogue box in which the planner can chose the needed functionality: a) general manipulation of sequences (creating, destroying and selecting sequences), b) visualisation and manipulation of modules assembly properties, d) automatic and manual sorting according to these properties to obtain an assembly sequence, d) visualising the assembly, e) motion planning for each module in the sequence and f) exporting the robotized cranes commands (see Figure 4). Once the assembly sequence and the robotized crane commands has been obtained, the assembly can be simulated in the AUTMOD3 simulation tool or the commands can be sent to the robotized crane for their execution. This crane is used to assembly modules in a demonstrator with 1:3 scale modules (see Figure 5). The AUTMOD3 software environment jointly with the robotized crane control system permits the automatic modular construction of the designed building in the demonstrator (Figure 6).

4. CONCLUSIONS

The AUTMOD3 planner, dedicated to obtain the assembly sequence and the motion planning for the automatic assembly of modules, has been presented. The key concepts that make application of this automatic planner feasible are exposed: a great integration of the assembly planning with the real assembly process has been achieved, and a suitable strategy of movement for modular assembly has been defined. This methodology allows to avoid algorithm complexity obtaining a fast and practical procedure. Finally, the automatic assembly of the modules employing AUTMOD3 tool and the robotized crane has been shown.

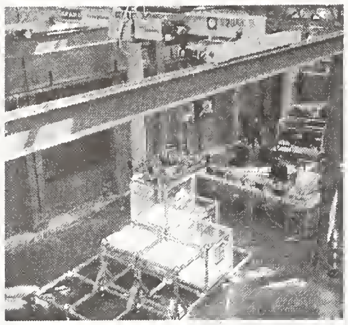


Figure 5. Carlos III University of Madrid robotized crane.

Figure 6. Robotized crane assembling a module.

5. REFERENCES

[1] Tracey, A., Child, T., Aouad, G., Brandon, P. and Rezgui Y., "Developing integrated applications for construction: the OSCON Approach", *I International Conference on Computing and Information technology for AEC,* Singapore, July, pp. 361-368, 1996.

[2] Boothroyd, G. and Alting, L., "Design for assembly and disassembly", *Annals CIRP*, 41(2), pp. 625-636, 1992.

[3] Canny, J., *The complexity of Robot Motion planning,* MIT press, 1998.

[4] Chen, P.C. and Hwang, Y. K., "SANDROS: A Dynamic Graph Search Algorithm for motion planning", *IEEE Transactions on Robotics and Automation*, 14(3), pp. 390-403, 1998.

[5] Diez, R., Abderrahim, M., Padrón, V.M., Celorrio, L., Pastor, J.M. and Balaguer, C. "AUTMOD3: An automatic 3D moularization system", 17th ISARC, Taipei, (Taiwan), pp. 1033-1038, 2000.

[6] Halperin, D., Latombe, J. C. and Wilson R.H., "A General Framework for Assembly Planning: The Motion Space Approach." *Algorithmica*, 26(3-4), pp. 577-601, 2000.

[7] Gupta, K., "Overview and State of the Art", *Practical Motion Planning in Robotics: Current Approaches and Future Directions*, Eds. Gupta, K. and del Pobil, P., John Wiley, pp. 1-8, 1998.

[8] Kavraki, L., Latombe, J-C. and Wilson R.H., "On the complexity of assembly partitioning", *Information Processing Letters*, 48(5), pp. 229-235, 1993.

[9] Kavraki, L., Latombe, J.-C., Motwani, R., and Raghavan, P., "Randomized Query Processing in Robot Path Planning", *Journal of Computer and System Sciences,* 57(1), pp. 50-60, 1998.

[10] Latombe, J. C., *Robot Motion Planning*, Kluwer Academic Publishers, 1991.

[11] Liu, Y-H. and Arimoto S., "Computation of the Tangent Graph of Polygonal Obstacles by Moving-Line Processing", *IEEE Trans. on Rob. and Automation*, 10(6), pp. 823-830, 1994.

[12] Overgaard, L., Larsen, R. and Jacobsen, N., "Industrial Applications of the AMROSE Motion Planner", *Practical Motion Planning in Robotics: Current Approaches and Future Directions*, Eds. Gupta, K. and del Pobil, P., John Wiley, pp. 155-183, 1998.

[13] Peñín, L.F., Balaguer, C., Pastor, J.M., Rodríguez, F.J. and Barrientos, A., "Robotized Spraying of Prefabricated Panels", *IEEE Robotics and Automation Magazine*, September, 18-29, 1998.

[14] Wolter, J.D., *On the automatic generation of plans for mechanical assembly*, PhD Thesis, The University of Michigan, 1988.

[15] Zha, X.F., Lin, S.Y.E. and Fok, S.C., "Integrated Intelligent Design and Assembly Planning: A survey", *Int. J. Adv. Manuf. Technol. 14*, pp. 664-685, 1998.

6. ACKNOWLEDGEMENT

This work has been supported by the EU FutureHome project (BRITE nº ER4-29671)

SIMULATION-FACILITATED FACTOR-BASED APPROACH FOR COST CORRELATION EVALUATION

Wei-Chih Wang

Associate Professor, Department of Civil Engineering
National Chiao Tung University, Hsin-Chu 300, Taiwan;
(886)3-5712121 ext. 54952; weichih@cc.nctu.edu.tw

ABSTRACT: Project cost becomes increasingly variable if many cost items for a construction project are correlated, and this can increase the uncertainty of completing a project within a target budget. This work presents a factor-based computer simulation model for evaluating project costs given correlations among cost items. Uncertainty in the total cost distribution of an item is transferred to several factor cost distributions according to qualitative estimates of the sensitivity of each cost item to each factor. Each cost distribution is then decomposed further into a family of distributions (children; costs given factor conditions), with each child corresponding to a factor condition. Correlations are retrieved by sampling from the child distributions with the same-condition for a given iteration of the simulation.

KEYWORDS: Computer Simulation; Cost Estimating; Risk Factor; Uncertainty

1. INTRODUCTION

Accurately estimating costs is an essential task in effectively managing construction projects. Each cost component, and thus project cost, is variable or probabilistic since future events are always uncertain [1]. Project cost becomes increasingly variable if several cost items are correlated, increasing the uncertainty of finishing a project to a target budget. Current research on correlated costs deals with theoretical issues concerning in the accuracy of correlations. For example, Touran and Wiser used a multivariate normal distribution to generate correlated cost variables for a precise simulation analysis, assuming that the correlation coefficients between variables are known [2]. The simulation model of Chau employed a percentile-based sampling procedure to influence the probability of sampling the same quantiles from two correlated probability density functions, according to whether the given correlation coefficient is positive or negative [3]. Finally, Ranasinghe highlighted some theoretical requirements, such as the conditions required to achieve a positive definite correlation matrix and the possibility of using an induced correlation to define the correlation between derived variables [4].

This paper presents a simulation-based cost model that considers correlations between cost items [5]. In contrast to existing cost related models in incorporating correlations, the proposed model is designed to meet the following three requirements which are considered practical in a cost management tool, namely: not requiring excessive input from management, introducing correlations indirectly (since this correlation information is not readily available) [2], and recognizing factor-based correlations when they occur in the field.

2. THE PROPOSED MODEL

2.1 Breakdown of uncertainty

The proposed model treats the cost of a bill item as a random variable. The cost variable is represented by a total cost distribution (that is, "grandparent" distribution) that combines a base cost with variations resulting from various factors. Variations owing to a particular factor are represented by a cost distribution, a "parent" distribution. The base cost is assumed to be deterministic, while the cost distribution for each factor is assumed to be a zero-mean random variable. Figure 1 schematically depicts this approach to break down the uncertainty. The base cost is taken to be the user's best estimate of an item's cost under expected factor conditions, and is the expected value of the total cost distribution for the item. Deviations from the expected value caused by various factors are introduced through the cost distributions.

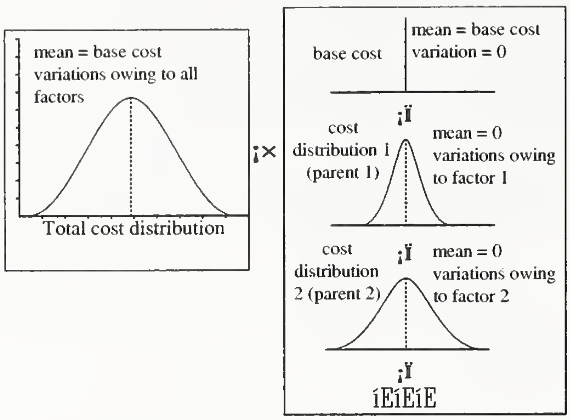


Figure 1. Breakdown of uncertainty

The model captures correlations by drawing cost samples from related portions of the cost distributions for cost items that are sensitive to a given factor. For example, the upper part of Fig. 2 classifies weather conditions into "better than expected," "normally expected," and "worse than expected." Based on these three different weather conditions, the weather related cost distribution is disaggregated into three corresponding child distributions (illustrated in the lower half of Fig. 2), namely, cost given better than expected weather (that is, better than expected weather child), cost given normally expected weather (that is, normally expected weather child), and cost given worse than expected weather (that is, worse than expected weather child). Child distributions may also overlap, as presented in Fig. 2. Restated, the cost of an item may be the same under both better than expected and normally expected weather conditions; or the cost with normally expected weather conditions may be less than the cost with better than expected weather.

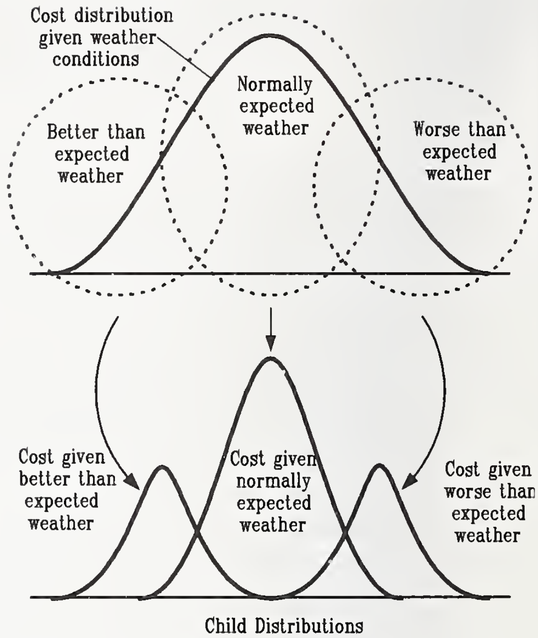


Figure 2. Decomposition of cost distribution into costs, given particular factor conditions

A cost model, in which the effect of uncertainty is broken down by factor, is derived from the unit cost perspective [5]. Following a series of derivation [5], C_i the cost of item i, may be expressed as

$$C_i = c_{i(0)} + \sum_{j=1}^{J} c_{i(j)} \tag{1}$$

where $c_{i(0)}$ is the estimated (or base) cost and the random variable $c_{i(j)}$, j = 1,..., J, is the cost (parent) distribution of cost item i due to factor j. Restated, Equation (1) displays the variations in the cost of an item, as a base cost and a series of cost distributions for various factors.

The model assumes that the costs of items are correlated only through the impact of shared factors. Different factors are assumed to cause independent effects. For example, assume that cost item 1 is sensitive to weather and labor, and cost item 2 is sensitive to weather and equipment. Only the weather-related cost distributions are correlated; the variations caused by labor and equipment are assumed to be independent. Then, regardless of the type of the marginal distribution of $c_{i(j)}$, the mean and variance of the cost of cost item i can be derived as [5]

$$M_i = m_{i(0)} + m_{i(1)} + m_{i(2)} + ... + m_{i(J)} = m_{i(0)} \tag{2}$$

$$\sigma_i^2 = SD_{i(0)}^2 + SD_{i(1)}^2 + SD_{i(2)}^2 + ... + SD_{i(J)}^2$$
$$= SD_{i(1)}^2 + SD_{i(2)}^2 + ... + SD_{i(J)}^2 \quad (3)$$

in which M_i and σ_i are the mean and standard deviation for C_i (the total cost distribution for item i), and $m_{i(j)}$ and $SD_{i(j)}$ are the mean and standard deviation for $c_{i(j)}$, with $SD_{i(0)} = 0$. The model finds M_i and σ_i for cost item i, and then determines $SD_{i(j)}$. In the example project presented herein, the three-point estimates of PERT are used to calculate M_i and σ_i.

In constructing a family of child distributions to represent changes in cost due to factor conditions, one goal is to preserve the mean and standard deviation of the cost distribution. In other words, the mean and standard deviation of the combination of the child distributions for a family should be the same as the mean and standard deviation of the cost distribution. Mathematically, this relationship can be represented [5]

$$m_{i(j)} = \sum_{h=1}^{H} p_{j(h)} \times o_{i[j(h)]} = 0 \quad (4)$$

$$SD_{i(j)}^2 = \sum_{h=1}^{H} p_{j(h)} \times (sd_{i[j(h)]}^2 + o_{i[j(h)]}^2) \quad (5)$$

in which H = number of child distributions; $p_{j(h)}$= probability of occurrence for child distribution h of factor j; and $o_{i[j(h)]}$ and $sd_{i[j(h)]}$ = mean and standard deviation, respectively, for child distribution h of factor j for cost item i. Equations (4) and (5) are valid for any type of statistical distribution. Steiner's theorem can be directly applied to justify (5) [6].

The mean of the child distribution for a given condition is the expected deviation from the mean of the cost distribution when the cost item is performed under the given condition. Means of child distributions are expressed through a variable x, the mean placement. The mean of each child distribution should be confined to a range that maintains the variance of the cost distribution. When x is equal to the limit, the child distributions will have zero standard deviations [5].

To construct a family of child distributions is to determine their means and standard deviations. Consider a cost distribution that is sensitive to factor j and has a variance of \$4 K. Assume that the user chooses the categories of better than expected, normally expected, and worse than expected conditions to describe the conditions of the factor. Then a family of three child distributions should be constructed. Assume that the probabilities of occurrence for the child distributions are equal; that is, $p_1 = p_2 = p_3 = 1/3$. Thus, based on (4) and (5), the mean and variance, respectively, of the combined child distributions are

$$\frac{1}{3}o_1 + \frac{1}{3}o_2 + \frac{1}{3}o_3 = 0 \quad (6)$$

$$\frac{1}{3}(sd_1^2 + o_1^2) + \frac{1}{3}(sd_2^2 + o_2^2) + \frac{1}{3}(sd_3^2 + o_3^2) = 4 \quad (7)$$

Assume $-o_1 = o_3 = x$ and $o_2 = 0$ so that (6) is satisfied, and let the child distributions have equal standard deviations, then (7) can be rewritten as

$$sd^2 + (2/3)x^2 = 4 \quad (8)$$

The limit of the value of x is found by requiring that the variance of the child distribution be non-negative. Namely,

$$sd^2 = 4 - (2/3)x^2 \geq 0 \quad (9)$$

Thus, the limit in this case is $x \leq \sqrt{6} = 2.45$ (limit = 2.45). In other words, the values of 2.45 and -2.45 are the two extreme means for Child Distributions 1 and 3, respectively. The next step is to select the value of x between 0 and 2.45. Instead of specifying the exact value of x, the proposed model suggests that the value of x be selected according to the level of influence of the factor under consideration on the cost item under consideration. In this example, assume x is set to one-half of the limit. Then x is equal to 1.27. The properties of this family of three child distributions are thus Child 1 ($p_1 = 1/3$, $o_1 = -1.27$, $sd_1 = 1.71$), Child 2 ($p_2 = 1/3$, $o_2 = 0$, $sd_2 = 1.71$), and Child 3 ($p_3 = 1/3$, $o_3 = 1.27$, $sd_3 = 1.71$).

2.2 Qualitative estimates

Cost distributions are derived according to subjective information. Project planners are asked to estimate qualitatively the extent to which each factor influences the cost of each item. For example, a cost item would be considered to be highly sensitivity to weather if its cost varies greatly depending on the weather. This approach of qualitative estimates is practical because the impact of uncertainties is easily expressed linguistically. No inherent restriction is placed on the number of levels of influence used for each factor. The example included herein use four levels of influence, high, medium, low, and no influence.

2.3 Scale system

A scale system is used to transfer the uncertainty associated with total cost distribution to the cost distributions based on qualitative estimates of the uncertainty sensitivity of cost item i to factor j [5][7]. That is,

$$\begin{aligned}\sigma_i^2 &= SD_{i(1)}^2 + SD_{i(2)}^2 + \ldots + SD_{i(J)}^2 \\ &= (w_1[Q_{i(1)}] + w_2[Q_{i(2)}] + \ldots + w_J[Q_{i(J)}]) \times K_i \\ &= (\sum_{j=1}^{J} w_j[Q_{i(j)}]) \times K_i\end{aligned} \tag{10}$$

$$SD_{i(j)}^2 = w_j[Q_{i(j)}] \times K_i \tag{11}$$

where $Q_{i(j)}$ is the qualitative estimate of the sensitivity of cost item i to factor j, and $w_j[Q_{i(j)}]$ is a scale for each level of influence. For example, the values of the estimates of *high*, *medium*, *low*, and *no* sensitivity for factor j can be represented by w_j[High], w_j[Medium], w_j[Low], and w_j[No], respectively. K_i is an adjustment constant that ensures that σ_i^2 is preserved. Since $w_j[Q_{i(j)}]$ is fixed for a given factor j, K_i will be different for each cost item. The value of w_j[No] is always zero. The value of $w_j[Q_{i(j)}]$ is higher when $Q_{i(j)}$ represents a higher level of influence. Consequently, a larger portion of the variance is distributed to a cost distribution that has a higher sensitivity.

2.4 Sensitivity of project cost to uncertainty

When several cost items for a project are sensitive to particular factors, these factors are likely to dominate the cost performance of the project. Knowledge of factor-sensitivities gives management a better idea of what factors to control. For instance, management should focus on carefully scheduling weather-sensitive tasks and ensuring adequate equipment is available if weather and equipment performance exert the biggest influence on project cost. Controlling the factors that influence performance improves performance more than modifying or changing work methods. This study measures the uncertainty sensitivity of each cost item to a given factor based on its standard deviation divided by its mean. A project in which a certain factor has a high standard deviation is considered highly sensitive to that factor (since the mean of project cost is equal for each factor), and consequently project cost is more likely to be affected by a change in that factor.

3. COMPUTER IMPLEMENTATION

In the model, when cost distributions are sensitive to the same factor, a sample cost is independently drawn from a particular child distribution (given a specified probability of occurrence) for each cost distribution. For example, if better than expected, normally expected, and worse than expected weather are equally likely to occur, then one-third of a predefined number simulation iterations will have cost samples that are simultaneously and independently drawn from the better than expected weather child distributions; one-third will have normally expected weather child distributions; and one-third will have worse than expected weather child distributions. A simulation language, STROBOSCOPE [8], is used to execute the simulation-relevant procedure described in the model. This procedure was implemented on a 586 PC with 64 MB under a 32-bit Windows environment (namely, Windows 98). Making 1,000 analyses of twenty-four cost categories of the example project took approximately six minutes, which is acceptable for research.

4. EXAPME DEMONSTRATION

An example for a building project is used to compare the results obtained using the model with two analyses that do not consider correlations, namely: a standard PERT analysis (PERT) and a Monte-Carlo simulation, carried out using normally distributed costs with the same mean and standard deviation as the model's total cost distribution (W/O Correlation Normal). Meanwhile, three different scale systems (Scales 1, 2, and 3) are applied to investigate the effect of the scale system. This project comprises 20 direct-cost division items and 4 indirect-cost division items (that is, insurance, tax, profit, and contingency). The model requires two types of inputs, the three-point cost estimates for each division item and the qualitative estimates of the sensitivity of each division item to various factors. The analyses considered here involve 1,000 simulation iterations. The scales of Scale 1 are listed in Table 1.

Table 1. Scales of Scale 1

	Scales			
F1	$w_{F1}[H]=16$	$w_{F1}[A]=12$	$w_{F1}[L]=8$	$w_{F1}[No]=0$
F2	$w_{F2}[Yes]=12$			$w_{F2}[No]=0$
F3	$w_{F3}[H]=7$	$w_{F3}[A]=5$	$w_{F3}[L]=3$	$w_{F3}[No]=0$
F4	$w_{F4}[H]=4$	$w_{F4}[A]=3$	$w_{F4}[L]=2$	$w_{F4}[No]=0$
F5	$w_{F5}[H]=3$	$w_{F5}[A]=2$	$w_{F5}[L]=1$	$w_{F5}[No]=0$

where "H", "A", "L", and "No" represent high, average, low, and no sensitivity, respectively. "Yes" or "No" are used to describe the sensitivity of cost items to F2. F1 - F5 represent owner approval, weather, material delivery, labor, and equipment, respectively.

Meanwhile, the scales for Scales 2 and 3 (which exaggerate the differences between high, medium, and low sensitivities) are displayed in Table 2 and Table 3, respectively.

Table 2. Scales of Scale 2

	Scales			
F1	$w_{F1}[H]=8$	$w_{F1}[A]=5$	$w_{F1}[L]=1$	$w_{F1}[No]=0$
F2	$w_{F2}[Yes]=8$			$w_{F2}[No]=0$
F3	$w_{F3}[H]=8$	$w_{F3}[A]=5$	$w_{F3}[L]=1$	$w_{F3}[No]=0$
F4	$w_{F4}[H]=8$	$w_{F4}[A]=5$	$w_{F4}[L]=1$	$w_{F4}[No]=0$
F5	$w_{F5}[H]=8$	$w_{F5}[A]=5$	$w_{F5}[L]=1$	$w_{F5}[No]=0$

Table 3. Scales of Scale 3

	Scales			
F1	$w_{F1}[H]=100$	$w_{F1}[A]=10$	$w_{F1}[L]=1$	$w_{F1}[No]=0$
F2	$w_{F2}[Yes]=100$			$w_{F2}[No]=0$
F3	$w_{F3}[H]=100$	$w_{F3}[A]=10$	$w_{F3}[L]=1$	$w_{F3}[No]=0$
F4	$w_{F4}[H]=100$	$w_{F4}[A]=10$	$w_{F4}[L]=1$	$w_{F4}[No]=0$
F5	$w_{F5}[H]=100$	$w_{F5}[A]=10$	$w_{F5}[L]=1$	$w_{F5}[No]=0$

Results: project cost. The project costs obtained from various analyses (PERT, W/O Correlation Normal, With Correlation Scale 1, Scale 2, and Scale 3) are compared using several metrics, namely the mean, standard deviation, minimum and maximum project costs. Table 4 lists the analytical results, and yields the following observations:

- The mean and standard deviations for PERT and W/O Correlation Normal are approximately the same because of the effect of the Central Limit Theorem.
- The analytical results with and without correlation analyses reveal very little difference in mean project cost. Restated, the correlation affects the variance rather than the expected cost.
- Correlation produces a project cost that may be significantly lower than expectations (e.g., \$117.96 K for Scale 1 versus \$132K for W/O Correlation Normal) or significantly higher than expected (e.g., \$184.25 K for Scale 1 versus \$167.49K for W/O Correlation Normal). The correlation effect thus has the potential to create an unexpected cost overrun.
- The project standard deviations of the three With Correlation analyses are 153%, 137%, and 149% higher than for the W/O Correlation Normal analysis for Scales 1, 2, and 3, respectively. For this example project, the choice of scale systems does not markedly affect the analytical results, which fact applies even in the case of Scale 3 (highlighting the differences between sensitivities), because the correlation effect determined by Scale 3 is enhanced only when most activities have high sensitivities to the same factor or factors. It was found out that the correlation effect tends to be dominated by the lower-sensitivity factor cost distributions, rather than the higher-sensitivity ones.

Results: uncertainty sensitivity. Table 5 summarizes the results of uncertainty sensitivity to F1, F2, F3, F4, F5, and all factors of project cost for different scale systems. For Scales 1 and 2, the project cost is most sensitive to F4 (labor), followed by F1, F5, F3, and F2. This information tells management that controlling the quality and availability of labor deserves special attention Meanwhile, in Scale

3, which increases the difference between high, medium, and low sensitivities, F1 becomes the most sensitive factor rather than F4. Notably, the PERT and W/O Correlation Normal models are unable to provide this type of sensitivity information.

Table 4. Comparisons of W/O correlation and the model analyses

Project Cost [a]	PERT	W/O Corr. Normal	Proposed model		
			Scale 1	Scale 2	Scale 3
Mean	150 [b]	149.97	150.06	149.69	150.33
Standard deviation	5.69	5.32	13.46	12.63	13.26
Min. cost	N/A	132	117.96	118.65	117.05
Maxi. cost	N/A	167.49	184.25	184.09	184.38

[a] The results are evaluated considering all factors.

[b] All data are expressed in thousands (K).

Table 5. Effect of scale systems

Factors	Standard deviation		
	Scale 1	Scale 2	Scale 3
1. Owner approval	7.9857[b][2][a]	6.8174 [2]	10.8664 [1]
2. Weather	1.5813[5]	1.9488 [5]	1.7313 [5]
3. Material delivery	2.7240[4]	1.9603 [4]	3.8811 [4]
4. Labor skills	8.3092[1]	8.9019 [1]	5.7696 [2]
5. Equipment breakdown	6.7841[3]	5.5836 [3]	4.0056 [3]
All factors	13.46	12.63	13.26

[a] [ ] indicates the rank of the sensitivity with respect to a given factor.

[b] All data are expressed in thousands (K).

5. CONCLUSIONS

This work has developed a simulation-facilitated factor-based model that allows correlation between cost items to be considered in cost analysis. The model is based upon the two-step breakdown of uncertainty. The correlation between cost distributions is caused by their sharing the same factor(s). Correlation is introduced by sampling from the child distribution representing a given factor condition. The use of qualitative estimates to describe the effect of factor-based uncertainty should make the user more comfortable in providing inputs than other approaches. Future research directions could include exploring ways to capture non-Normal cost distributions and total cost distributions; implementing time-dependent and non-time-correlated cost variables; and applying the proposed model to other practical projects.

6. ACKNOWLEDGEMENTS

The author thanks Professor Laura Demsetz from the College of San Mateo and Professor Hojjat Adeli from the Ohio State University for their valuable assistance. Dr. Julio Martinez is also commended for making STROBOSCOPE available.

7. REFERRENCES

[1] Adeli, H. and Wu, M., "Regularization Neural Network For Construction Cost Estimation", Journal of Construction Engineering and Management, ASCE, Vol. 124, No. 1, pp. 18-24, 1998.

[2] Touran, A. and Wiser, E.D., "Monte Carlo Technique With Correlated Random Variables", Journal of Construction Engineering and Management, ASCE, Vol. 118, No. 2, pp. 258-272, 1992.

[3] Chau, K.W., "Monte Carlo Simulation Of Construction Costs Using Subjective Data", Construction Management and Economics, Vol. 13, pp. 369-383, 1995.

[4] Ranasinghe, M., "Impact Of Correlation And Induced Correlation On The Estimation Of Project Cost Of Buildings", Construction Management and Economics, 18, pp. 395-406, 2000.

[5] Wang, W-C., "Simulation-Facilitated Model For Assessing Cost Correlations", Journal of Computer-Aided Civil and Infrastructure Engineering, Vol. 17:5, pp. 368-380, 2002.

[6] Levin, R.I. and Rubin, D.S., Statistics for Management, 5th Edition, Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[7] Wang, W-C. and Demsetz L. A., "Model For Evaluating Networks Under Correlated Uncertainty – NETCOR", Journal of Construction Engineering and Management, ASCE, 126(6), pp. 458-466, 2000.

[8] Martinez, J. C., STROBOSCOPE: State and Resource Based Simulation of Construction Processes, Ph.D. Dissertation, University of Michigan, Ann Arbor, Michigan, 1996.

CRITERIA OF TAKING DECISIONS AT TECHNICAL OPERATION OF BUILDINGMACHINERY IN CONDITIONS OF THE CONSTRUCTION PLANT

Alexej Bulgakow*, Alexander Drownikow, Viktor Perschin****

**Technical University Munich, a.bulgakow@gmx.de*
*** South Russia State Technical University*, drovnik@sssu.ru

Abstract: In the article on the basis of similarity functioning the technique of getting and using the criteria of taking decisions at technical operation control, modernization and replacement of the elements of building machinery in conditions of the construction plant is considered.

Keywords: similarity, functioning, criteria, technical operation, and modernization.

INTRODUCTION

The issue of quality management of technical operation (at maintenance service, repair, and also modernization, technical state control, solving problems of refused elements, units replacement, etc.) of building machinery in conditions of the construction plant remains quite acute, inspite of positive experience of its solving in stationary conditions.

Methods and means of solving the problems mentioned above available and used in the world practice are only suitable for conditions of stationary specialized repair - diagnostic firms and building industry enterprises. Therefore at practical solving of the specified problems in conditions of the construction plant they are guided by the available operational experience and the employees' opportunities. Such kind of methods (tests and mistakes) at the building machinery diagnostics, analogues selection instead of the replaced or refused original elements and systems at modernization obviously results in reduction of operational and technological opportunities of the machinery as a whole. It basically happens because of the discrepancy of constructive and functional parameters of the compared analogues. Therefore a machine or a unit will be represented as though another technical system that should be adjusted in accordance with the original one some additional technical or technological measures (adjustment, a mode change, etc.) are needed.

Engineering methods of taking decisions in the specified conditions should be based on the system and, if possible, include typical techniques of calculations, procedures of taking decisions. As a result of engineering calculations criteria of the discrepancy allowable, criteria of similarity compared (initial and received) systems and a way of preset values achievement of the discrepancy should be received.

For solving of the specified problems it is possible to use principles of similarity of system's functioning representing one of new kinds and methods of research in the classical theory of similarity. This theory states that any functional dependence between physical parameters of the researched phenomenon and an object can be also expressed as dependences among the criteria of similarity made of these parameters. The features of criteria are the following: informatively, models of refusals development in the criteria form typical for similar processes of various technological systems; complex criteria of similarity - independent parameters. All these features allow to simplify models of development of machinery refusals and also to use algorithms of evaluation of its reliability, conformity which allow recalculation with their help of the initial or operational machinery parameters from one conditions to the others.

The procedure of estimation of the elements of the machine, the degree of the adequacy of functional characteristics and the physical processes carried out at various stages of the physical and moral obsolescence, and also the order of elimination of their possible mismatch runs as following:

- To reveal in and out parameters of the technological process carried out by a system or a machine;
- To reveal determining parameters of the carried out physical process and the parameters describing the work capacity and design for the analyzed element;
- To form criteria of similarity of functioning describing the constructional features and the physical processes carried out for the research element, a subsystem; containing it and a machine as a whole;
- To establish the dependence of criteria of the machine on the dependence of the elements, subsystems;
- To define the numerical values of similarity criteria at the initial and the researched moments of operation (for example, before and after the replacement of an element with the analogue), and also the accuracy of

reproduction of the given technological process;

- To take decisions according to the specific purpose at the insufficient accuracy of the reproduction of the preset values of the parameters of the technological process by an element on the similarity criteria of functioning.

Let us admit that the "out" physical parameters of the technological process carried out by the machine, are y_1, y_2, … , y_n and the "in" specifying and revolting influences, the physical process and the constructional features of the element are determined by parameters x_1, x_2, … ,x_m, and at least one of y_i parameters is connected with x_i, i.e.

$$y_i \in \{y\} = f(x_1, x_2, \ldots, x_m). \quad (1)$$

The choice of those or other parameters x_i is defined by a specific task of research of the technical system and by the following taking of decisions.

Let us accidentally admit that the quantity of parameters x_i is equal to 7 and they are characterized by four units of measurements: time T, weight M, length L and temperature θ. According to the theory of similarity and dimensions, the number of criteria of similarity will be equal 3 (the general number of parameters minus the number of the independent dimensions). These criteria will be defined by the method of zero dimensions. Let us conventionally take x_3, x_4, x_6, x_7 as the key parameters. The determinant made for these parameters from parameters of degrees of their dimensions, should be different from zero, i.e.:

$$\Delta = \begin{vmatrix} \mu_3 & \lambda_3 & \tau_3 & \eta_3 \\ \mu_4 & \lambda_4 & \tau_4 & \eta_4 \\ \mu_6 & \lambda_6 & \tau_6 & \eta_6 \\ \mu_7 & \lambda_7 & \tau_7 & \eta_7 \end{vmatrix} \neq 0 \quad (2)$$

This term means, that taking parameters are really independent. Symbols $\mu_i, \lambda_i, \tau_i, \eta_i$ in the determinant (2) correspond to the exponents at dimensions of weight M, lengths L, time T and temperatures □ for the taken parameters x_j.

As the criteria of similarity represent sizes of zero dimension the expression (1) should by written down in the dimensionless way:

$$\frac{y_i}{x_3^{\alpha_y} x_4^{\beta_y} x_6^{\gamma_y} x_7^{\delta_y}} = \Phi\left(\frac{x_1}{x_3^{\alpha_1} x_4^{\beta_1} x_6^{\gamma_1} x_7^{\delta_1}}, \frac{x_2}{x_3^{\alpha_2} x_4^{\beta_2} x_6^{\gamma_2} x_7^{\delta_2}}, 1, 1, \frac{x_5}{x_3^{\alpha_5} x_4^{\beta_5} x_6^{\gamma_5} x_7^{\delta_5}}, 1, 1 \right) \quad (3)$$

where dimensions of the denominators and numerators should be equal.

Substituting the complexes of expression (3) with the dimensions of the appropriate parameters x_j, by known from the theory of similarity methods which numerical values are designate accordingly α_j, b_j, c_j, d_j.

The criteria of similarity will finally become:

$$\pi_1 = \frac{x_1}{x_3^{a_1} x_4^{b_1} x_6^{c_1} x_7^{d_1}};$$

$$\pi_2 = \frac{x_2}{x_3^{a_2} x_4^{b_2} x_6^{c_2} x_7^{d_2}};$$

$$\pi_3 = \frac{x_5}{x_3^{a_5} x_4^{b_5} x_6^{c_5} x_7^{d_5}}; \quad (4)$$

$$\pi_{y_i} = \frac{y_i}{x_3^{a_y} x_4^{b_y} x_6^{c_y} x_7^{d_y}}.$$

Depending on the physical meaning of the received criteria of parameters included the separate criteria will characterize geometrical similarity of the elements - analogues, others - similarity of physical processes occurring in them. Accordingly the first ones should be used for selection of elements - analogues at their constructive replacement of the refused elements, and others - for correcting of the technological parameters carried out by the machine. Thus it means, that the received criteria of similarity can be accidentally divided or multiplied thus forming new criteria of similarity suitable for solving of specific tasks. In particular, it is possible to form the special criteria for y_i out parameters of systems' functioning. These criteria include "in" influences and internal parameters of the systems, the so-called criteria of similarity of elements functioning, the systems of building machinery.

Generally criteria may have an area of the admitted values as constructional or the technological parameters of the element of the machine are given with the accuracy in certain limits. Therefore values of criteria i of the element - analogue π_{ia} and element - original π_{i0} should follow the term:

$$\pi_{ia} = \pi_{i0} \tag{5}$$

where the set of values π_{i0} is restricted by bottom $\underline{\pi_{i0}}$ and upper $\overline{\pi_{i0}}$ limits.

According to the given technique the concrete examples of taking decisions at technical operation of the technical systems should be considered.

Thus if the analysis of interchangeability at a choice of analogue the term is broken, recalculation and the change of the rating values of the parameters carried out by the element of the technological process with the purpose of saving the value of the main, determining quality criterion of system functioning or the machine as a whole made or calculated from the condition of the given productivity, effort or other parameters is necessary. If quality of functioning is not saved after that it means that another element of the replaced element is needed. The order of taking decisions about the interchangeability of the technical objects was considered by the replacement of the refused pump adjustable on the pressure such as PVE 35 QR-1-21-CVP-JT6 by the domestic analogue.

The pump serves for maintenance of supply of the working element, carrying out of the working operation in time $t_0 = 2 \div 2{,}5$ sec. that corresponds to the oil submission into the cylindrical cavity Q_c = 70-80 l/min, and also for the shneck rotation with frequency $n_s = 120$ min^{-1} that corresponds to the oil expenditure through hydro motor Q_{gm} = 55 l/min at pressure P = 12 MPa. The import pump is characterized by the following rating values of parameters: submission Q_n = 106 l/min; working volume q_n = 70 sm^3/vol; pressure P_n = 25 MPa; rotation frequency of the shaft n_n = 1450 min^{-1}; the maximal pressure of forcing P_{nmax} = 35 MPa. As a working liquid oil viscosity ν_n = 46 mm^2/s.

The criteria of similarity made in accordance with the offered technique for the given conditions look like:

$$\pi_1 = \frac{t_0 Q_n}{q_n} = 193{,}4$$

$$\pi_2 = \frac{n_{gm} q_n}{Q_n} = 0{,}024$$

$$\pi_3 = \frac{P_n}{P_{n\max}} = 0{,}75; \tag{6}$$

$$\pi_4 = \nu\sqrt{\frac{F}{Q_c}} = 39{,}7$$

with F as the area of the piston.

Apparently from the first and the second criteria dependences, saving of preset values of technological parameters t_0 and n_{gm} at replacement of the pump by its domestic analogue is possible if at π_i = const the condition is satisfied:

$$\frac{Q_n}{q_n} = \frac{Q_{n.a}}{q_{n.a}} = \text{const.} \tag{7}$$

However, taking into account, that the submission of oil in the hydro system of the machine to the hydro motor and the cylinder is adjusted and stabilized by the stream regulators, this condition basically may not be followed and the estimation of the interchangeability of the pumps should only be carried out on conformity of rating values Q and P.

At the same time at an estimation of the interchangeability of the pumps adjustable by the pressure of the pumps the third criteria is of great importance as it means, finally, steepness of the static characteristic of the adjustable pumps, the time and accuracy of regulation.

The carried out analysis on the criteria of the interchangeability of the import pump by the domestic it is allowed to choose this analogue. In this case as an analogue the domestic adjustable pump RNA 1D 63/320 UHL4 with the following parameters was taken: $Q_{n.a}$ = 87 l/min; $q_{n.a}$ = 63 sm^3/vol; $P_{n.a}$ = 32 MPa; $P_{n.a.max}$ = 40 MPa; $n_{n.a}$ = 1500 min^{-1}. The oil viscosity recommended by the factory manufacturer is $\nu_{n.a}$ = 12□75 mm^2/s. According to this recommendation and including the fourth criterial dependences (6) oil Tp-46 GOST R 9972-74 was agreed. The value of the criterion π_3 = 0,8, is better for the given pump than at the replaced one.

Similarly an example of taking decisions at the analysis of the technical condition of building machinery should be considered. The choice and the use of these or those techniques and means of the technical condition of machinery in terms of the concrete building enterprises depends on the following factors: constructive and repair complexity of building machinery, their quantity at the construction plant, economic feasibility of application of the techniques and means of diagnostics, their complexity and availability. Thus the engineering preparation of the testing process of the technical state is of great importance including the determining of "in", internal and tested parameters and also their boundary condition at an estimation of the work capacity of the machinery; the grounded choice of kinds and evaluative criteria of the technical conditions and the appropriate to the evaluations recommendations on the regulations of the further operation the technical service and repair of the concrete machines.

Criteria of similarity of functioning having informatively according to the processes carried out by the machine or the mechanism obviously may be the source of information concerning the way how far the concrete technical condition corresponds with the efficient or another given condition. And if the number of the discrepancy of the initial and valid values of the criteria exceeds the preset limiting value it is possible to conclude about the disability of the machine.

The advantage of the criteria of similarity is also in the fact that they can be made of any process parameters, the phenomena connected by the functional dependence. At the control of the technical condition the criteria dependences made of the structural and diagnostic parameters are also of interest. In this case it is possible to determine the value of the structural and diagnostic parameter (which registration is complicated) using accessible techniques and means of registration of the diagnostic parameters be the criteria dependence without operations of dismantlement and without breaking the manufacturing process in which the machine is involved.

The control technique of the technical condition of the element, subsystem of the building machine consists of the following.

Accepting values of the parameters included into the expression (4) and equal to the values appropriate to the given condition of the element the initial values of the criteria of similarity are determined.

In the process the control of technical conditions having defined (by calculating) the valid values of diagnostic parameters y_i and having accepted constant value criterion i (π_i = const), according to the expressions (4) it is possible to find the valid value of structural parameter X_{ij} On comparison of the parameter value with its boundary (the greatest $\overline{X}_{ij}$ and the least $\underline{X}_{ij}$) values appropriate to the given condition of the element it is possible to judge about the valid technical condition of the element and the machine as a whole, i.e., if

$$X_{ij} \in \left| \overline{X}_{ij}, \underline{X}_{ij} \right|,$$

element j on parameter X_{ij} is in the efficient condition or on the contrary if

$$X_{ij} \notin \left| \overline{X}_{ij}, \underline{X}_{ij} \right|$$

the valid condition does not correspond to the preset value, i.e. the element is disable.

The presented technique was tested at the evaluation of the technical condition of the hydro drive. Thus the following parameters of working capacity were considered: δ - a backlash in the interface of the pair "the piston and the cylinder"; μ - the oil viscosity; Q_0 – submission of the pump of the constant productivity.

Let us admit that the functional dependence determining the work of the hydro drive establishes the connection of time t of the turn of the working element with pressure difference on the cylinder of turn ΔP_c, viscosity μ and the module of oil elasticity E, the size of the outflow in cylinder Q_y through the backlash δ in condensation, diameter d, length of course H and area F of the piston.

The criteria of similarity of the preset process of functioning were received, the main of them are:

$$\begin{aligned} \pi_1 &= t\, Q_0\, \delta^{-3}; \\ \pi_2 &= \Delta P_c\, \delta^3\, \mu^{-1}\, Q_0^{-1}; \\ \pi_3 &= t\, \mu^{-1}\, \Delta P_c. \end{aligned} \tag{8}$$

As measured parameters it is convenient to accept t and ΔP_c, and as seeked parameters - δ, μ, Q_0.

According to the first and third criteria of similarity of functioning the increase of the turn time can be generally connected with the appropriate decrease of the pump submission, decrease the viscosity and the increase of the backlash of the pressure difference on the cylinder.

According to the second criterion the change of the pressure difference on the cylinder (at constant loading) may appear at the changing of the backlash, the oil viscosity and the submission of the pump.

The strategy of the control of the technical condition on the criteria of similarity can be different according to the concrete conditions of operation and construction of the hydro drive but the essence of the technique remains constant. The following strategy can be used in this case:

1. The initial values of the parameters included into the criteria dependences and appropriate to the preset condition were considered:

t = (5 – 6), s;
$Q_0 = (5{,}0 - 5{,}0)10^{-4}$, m^3s^{-1};
$\delta = (1{,}5 - 2{,}0)10^{-5}$, m;
$\Delta P_c = (4{,}0 - 2{,}5)10^{+6}$, $N \cdot m^{-2}$;

$\mu = (182 - 160), N \cdot s \cdot m^{-2}$,

where the first ones are optimum, and the second ones - the limiting values of the parameters satisfying the preset, available and efficient condition of the hydro drive.
2. Values of criteria of similarity paid off:

$\pi_1 = 86 \cdot 10^{10}$; $\pi_2 = 12{,}75 \cdot 10^{-8}$; $\pi_3 = 10{,}95 \cdot 10^4$.

3. The valid values of the time of the turn and the pressure difference were measured:

$$t_u = 8\ s;\ \Delta P_{cn} = 3 \cdot 10^6, N \cdot m^{-2}.$$

Taking into account the real conditions an possibility of calculation of these or those parameters in case of their preserving at the level of initial values of the other parameters was researched. The calculation of the parameters was carried out.

Using the third criteria dependence and having assumed that the value of the backlash in the condensation and the quantity of the pump submission corresponds with the preset the value of the oil viscosity change was determined as:

$$\Delta\mu = \frac{t_u\, p_{cu}}{\pi_3} - \mu\ , \qquad (9)$$

and its calculating value accepted for the valid one.

The calculating value of the oil viscosity $\left(\mu_p = 146,\ N \cdot s \cdot m^{-2}\right)$ is significantly less of the required one that allows to make a conclusion about the necessity of its analysis and replacement. Thus the real submission value is determined at the real viscosity value, the measured pressure difference value and the initial value of the backlash. Similarly, the correctness of the assumption of the backlash conformity is checked.
In this case the values are within the limits and make: the pump submission $Q_r = 5{,}4 \cdot 10^{-4}$, m^3. s^{-1}; the backlash in condensation $\delta_r = 1{,}7 \cdot 10^{-5}$, m. It accounts for the working condition of the pump and the cylinder, and the condition of the hydro drive is disable according to "the oil viscosity" parameter.

The presence of the functional dependences in the form of the criteria of similarity of conditions after measuring of the true values of the diagnostic parameters of the elements allows to calculate the real values of their structural or some other diagnostic parameters and as a result to determine the technical condition of the tested element, the system or the machine, the unit as a whole. And this in its turn helps to form the strategy of their further operation (adjustment, repair; replacement of defective elements) without equipment downtimes necessary for these purposes earlier and without any dismantlement works.
Nowadays the caring out of the offered technique of testing of the technical conditions is quite feasible by the experts - mechanics of the building organizations.

The similar procedures of taking decisions can also be carried out at modernization of the building machinery.

It is necessary to note, that there are some ways of getting of the criteria of similarity on the basis of the preset nomenclature of the parameters characterizing the physical essence of the researched process. However the most effective is the way allowing to use the computer facilities. For this purpose some simple computer programmers are developed by the authors. The received criteria of taking of the decisions at the technical operation are typical and universal for use at the various stages of the "life cycle" of the certain class of the building machinery.

CONCLUSION

Criteria of similarity of functioning having informatively according to the processes carried out by the machine or the mechanism obviously may be the source of information concerning the way how far the concrete technical condition corresponds with the efficient or another given condition. If the number of the discrepancy of the initial and valid values of the criteria exceeds the preset limiting value it is possible to conclude about the disability of the machine.

The advantage of the criteria of similarity is also in the fact that they can be made of any process parameters, the phenomena connected by the functional dependence. At the control of the technical condition the criteria dependences made of the structural and diagnostic parameters are also of interest. In this case it is possible to determine the value of the structural and diagnostic parameter (which registration is complicated) using accessible techniques and means of registration of the diagnostic parameters be the criteria dependence without operations of dismantlement and without breaking the manufacturing process in which the machine is involved.

DATABASE SYNCHRONIZATION TECHNOLOGY FOR MULTI-PROJECT SCHEDULE COORDINATION

by

Hyun Jeong Choo[1] and Iris D. Tommelein[2]

ABSTRACT

This paper describes a computing environment, called WorkMovePlan, that supports the exchange of data pertaining to resources shared between multiple production units and between multiple projects. It presents issues related to the generation and management of this data exchange and the development of a distributed, multi-project scheduling system that is deployed in industry practice.

KEYWORDS: Planning; Scheduling; Distributed Scheduling; Multi-project Scheduling; Database; Synchronization; WorkMovePlan.

1. INTRODUCTION

In a complex and dynamic construction project, no single participant can work in isolation for long. In addition, many participate in several projects at the same time. Work of every participant is interwoven with work of others. This is especially true for those responsible for production—designers, construction personnel, and other specialists who as individuals or as a team make up a production unit (PU)—as their deliverables are prerequisites to the work of others.

Production activities of PUs are interlinked because of physical dependencies and resource dependencies (here, resources are information, material, personnel, equipment, and space). Whereas physical dependencies clearly determine activity sequencing (e.g., in-wall electrical and plumbing systems have to be placed before wall panels are installed), resource dependencies do not: multiple activity-sequencing alternatives might exist.

From the perspective of the PU performing these activities, one alternative might not have a clear advantage over another, but from the perspective of others, or from the perspective of the project as a whole, it may be superior, for instance, if it "releases" more work or more resources. Conversely, what one PU identifies as a superior alternative may be inferior on a broader systems basis.

The WorkMovePlan system described in this paper provides database and graphical support for PUs to explore and rank alternatives, but it keeps people in the loop; WorkMovePlan does not automate this process. Job-shop scheduling, multi-objective decision-making, and heuristic optimization are needed to gauge and trade off what is best for individual PUs vs. the system, but discussing these is beyond the scope of this paper.

2. RESOURCE- vs. PROJECT-CENTRIC DATA

Planners can assess the value of and compare alternatives only if and when activity descriptions are detailed enough and especially when shared resource assignments are made sufficiently explicit (exactly how detailed and specific they have to be depends on the situation). However, a single planner probably does not have the knowledge to provide all that is needed. PUs have to develop their own

[1] PhD Candidate, Construction Engineering and Management Program, University of California, Berkeley, 215 McLaughlin Hall, Berkeley, CA 94720-1712, choohj@ce.Berkeley.edu

[2] Professor, Construction Engineering and Management Program, University of California, Berkeley, 215-A McLaughlin Hall, Berkeley, CA 94720-1712, tommelein@ce.Berkeley.edu, TEL 510/643-8678, FAX 510/643-8919

resource-loaded schedules based on each member's skill level and productivity and then share them with others to allow for comparison and the selection of alternatives that best suit multiple interests.

To make effective trade-offs and generate realistic schedules, planners need to adopt a more resource-centric view than they traditionally have. Resource-centric means that the schedule for each resource is the building block for the project and resources may be engaged on multiple projects. This view is contrary to the view adopted by many web-based project management systems on the market today, which are project-centric. Other researchers also adopt a multi-project view, as is presented here. For instance, Scherer et al. (2002) propose that each project participant plan their workflow within the framework of all projects s/he is involved in.

A PU's resource-loaded schedule may contain information about two types of resources: (1) dedicated resources and (2) shared resources. Dedicated resources are committed solely to a single PU on a single project. Shared resources are committed to more than one PU or to more than one project. Some shared resources may serve multiple PUs on multiple projects, which complicates the coordination problem even more. Shared resources may be project shared or company shared.

Project-shared-resources are resources used by several PUs, but not necessarily 'owned' by any. Examples are material hoists used by any or all on site, but also personnel such as project management staff (project engineers and superintendents), and space such as material storage areas, pre-installation and installation working areas, and access paths. Company-shared-resources belong to a specific owner who is engaged in several projects. Examples are equipment such as expensive hoisting equipment, large plotters, and personnel such as project managers and safety inspectors. This distinction affects who has a say and what objectives must be met in selecting from alternative resource assignments.

Information regarding which projects share a resource is tracked by the resource's owner or whomever obtained (e.g., rented) the resource, whereas information regarding which PUs share a resource resides in a production schedule of each PU. In order to ensure that assignments do not result in conflict, resource allocation needs to be checked across multiple projects and PUs.

Several difficulties exist in achieving this goal using current project (and production) management tools: (1) each PU must develop an appropriately-detailed resource-loaded schedule and make a significant amount of tacit planning knowledge explicit, (2) these schedules must be described in a common language so that they can be understood by others, (3) data for these schedules must be maintained in each company's database, while schedules are being coordinated, alternatives negotiated, and conflicts resolved, (4) data must be reliable and disseminated in a timely fashion.

3. WORKMOVEPLAN

3.1 Design Objectives

WorkMovePlan (Choo and Tommelein 2000a, 2000b), a computing environment designed to support distributed planning, allows each PU's planner to create their own schedule. WorkMovePlan's aim is to help project participants create more reliable schedules, in an effort to make the project delivery process more lean (also see www.leanconstruction.org).

The WorkMovePlan environment builds on Microsoft Access (Microsoft 2000a) and Microsoft Visio (Microsoft 2000d). Various forms in Access allow the user to input and manage a detailed activity list and resource assignments based on the Last Planner methodology (Ballard and Howell 1994, Choo et al. 1999). These activities can be directly imported from Microsoft Project (Microsoft 2000c) or they can be developed from scratch. The hierarchical structure of the activities allows the user to break them down to any level of detail. Since WorkMovePlan does not rely on a single person but instead relies on any or all production managers (superintendents and foremen) to enter information, the description of activities and resource assignments can get very specific. The link to Visio allows planners to geometrically detail space use based on space layout stencils.

WorkMovePlan captures data pertaining to multiple projects. A project-specific detailed schedule is automatically shared with other project participants using database synchronization technology. The PUs can then check for shared resource conflicts within each project and across projects.

WorkMovePlan maintains an offline copy of other PUs' schedules as well as its own. It automatically updates changes only upon synchronization. WorkMovePlan is portable and does not require a fast consistent Internet connection. It can thus be used by practitioners, including even those who do not have consistent Internet access to interact in real time with an online database.

Each planner can look at the detailed production schedule including space use on site (described later) for all project participants and determine whether they result in conflict. The planners then need to collaborate off-line with others to develop alternatives for specific conflicts and determine which alternative best meets the needs of those in conflict, of the projects they are involved in, and of their companies.

3.2 'Near real time' Data Sharing based on Synchronization

WorkMovePlan's distributed planning and coordination feature relies on 'near real time' data sharing, which is based on the technology called synchronization. Synchronization is defined as "the process of updating two replicas in which all updated records and objects are exchanged. The exchange of data between two replicas can be one-way or two-way" (Microsoft 1999).

Each WorkMovePlan is a replica that is two-way synchronized. Each replica's database contains two parts: one that contains private information and another that contains public information. Private information concerns the owner of the database. It contains information regarding its resources, associate costs, and the detailed schedule of each PU. This information is not exclusive to a single project since resources may be shared across multiple projects. The WorkMovePlan user can thus schedule multiple projects at the same time.

WorkMovePlan automatically generates public information by filtering out what is unnecessary to share, based on pre-set conditions. For example, Figure 1 shows the screenshots from WorkMovePlan for roof drain installation. The bottom portion shows the detailed weekly work plan for contractor 'Atlantic Roofs' as seen by its employees. The top portion shows the weekly work plan for Atlantic Roofs as seen by all other project participants. Accordingly, the private information names Gilbert Atlas as the PU and the exact hours (4.5, 8, and 5.5) he is scheduled to work. In contrast, the automatically-generated public information shows the name of the company the PU belongs to (Atlantic Roofs) and only the days (Wednesday, Thursday, and Friday) when work will be done. By making a commitment at a less detailed level to other project participants, the PU creates flexibility to carry out work within any part of the revealed duration. This is satisfactory as long as the output thus delivered does not to prevent others from performing their work.

Public information is the replicated part of WorkMovePlan (Figure 2). By replicating public information between all replicas of WorkMovePlan, schedule information regarding all PUs can be automatically updated. A similar data categorization is used in Microsoft Exchange Server (Microsoft 2000b), which can be configured to contain private- as well as shared information. The shared information can be created and viewed by any one who has been granted permission to do so. However, private information is accessible only by each so-designated individual and not by anyone else.

WorkMovePlan automatically synchronizes only information that is relevant to each project (Figure 3). The main reason for designing the database in such a way, rather than using a centralized on-line database, was that not all PUs have a consistent Internet connection. Despite recent advancements in information technology, project site offices rarely have high-speed Internet access, especially at the start of a project. Project managers have pointed out that their planning system has to be in place from day one (at the latest!) because once the project starts, it is very hard for them to learn and/or change procedures and support tools. Another reason is that many PUs are protected by company- and project-specific firewalls. These firewalls, in many cases, prevent users from taking advantage of the available full speed of their Internet connection. Should online planning tools be used during meetings, progress of the meeting would slow down to match the Internet connection speed.

A disadvantage of using synchronization technology is that data is not available to all project participants in real time. Project participants might not synchronize WorkMovePlan for some time, but still create their own plans based on obsolete data from others. This may result in conflict between project participants' schedules and create rework when synchronization takes place.

By keeping a copy of the 'near real time' information, i.e., the information that was available the last time the database was synchronized, the owner of each WorkMovePlan replica can still view the schedule information of others off-line.

4. EXTENDED RESOURCE PLANNING

WorkMovePlan extends planning to include space scheduling (Tommelein and Zouein 1993). A planner can specify site space needs on a day-to-day basis for labor, equipment, and materials in terms of work-, laydown-, staging area, or access path as needed throughout the execution of a work package, which is the unit of work assigned to a PU. WorkMovePlan requires the user to explicitly input information on resources that need to be considered during space scheduling (Figure 4). This space scheduling information is automatically synchronized in the same way as is done for other resources.

Default categories for space scheduling refer to material, equipment, and labor but others can be included as needed. Shape refers to the physical shape of the space required. X, Y, and Z refer to the dimensions of the needed space. Although three dimensions are specified, WorkMovePlan's space scheduling takes place in a 2-D environment.

2-D drawings (such as blueprints showing a site arrangement or a building floor) are widely available and space can be assigned easily in 2-D. 2-D layouts convey space scheduling information in a straightforward fashion. They are crude but adequate for this application. Nevertheless, the height dimension entered by the user can later be combined by WorkMovePlan with the layout schematic to generate a 3-D virtual reality mock-up using the Virtual Reality Modeling Language (VRML 1995). Figure 6 shows a sample VRML model that is automatically generated from WorkMovePlan.

The default schedule for space use is from the start- to the end date of the work package, but it can be adjusted to represent other realities, such as the delivery of materials a day prior to the start of the work package. Once all resources to be assigned are specified, their positions can be selected using a graphical user interface (GUI). WorkMovePlan builds on Microsoft Visio as the GUI for space scheduling. All information generated within Visio is captured by WorkMovePlan and shared across all project participants. Planners can view other participants' space use when scheduling their own space use. Choo and Tommelein (1999) describe an example application of WorkMovePlan.

5. CONCLUSIONS

The ability to build realistic schedules for projects as well as for individual PUs depends heavily on being able to collect and distribute reliable information. The most reliable information regarding resource characteristics (e.g., productivity and availability) resides with each PU. However, having information from each PU does not necessarily guarantee a realistic schedule unless the planning process itself promotes realistic planning. The realism of schedules also depends heavily on timeliness of the data being used. Each participant has to create and provide data to other participants with sufficient lead time to allow for conflict detection and resolution.

WorkMovePlan is a tool that helps to collect and capture such data, and it makes selected data available for sharing with other project participants. WorkMovePlan's ability to make detailed assignments in terms of labor, equipment, and space will allow project participants to generate more realistic schedules than they currently do.

WorkMovePlan suggests a very different way of coordinating project participants as compared to what is done in current practice, which includes numerous 'throw-away schedules' that so many PUs generate today (Russell and Froese 1997). Resistance is expected when a new planning paradigm is presented. Choo and Tommelein (2001) discuss several barriers to adoption in industry practice of the Last Planner methodology and the WorkMovePlan environment. It remains to be seen whether the industry will widely embrace either one or both. In the mean time, additional research is to result in better tools for job-shop scheduling, multi-objective decision-making, and heuristic optimization, which can then be integrated with WorkMovePlan.

6. REFERENCES

Ballard, G. and Howell, G. (1994). "Implementing lean construction: stabilizing work flow." *Proc. 2nd Ann. Conf. on Lean Constr.*, Santiago, Chili, Sept. Reprinted in Alarcon, L. (1997)(ed). *Lean Construction*. A.A. Balkema, Rotterdam, The Netherlands, 497 p.

Choo, H.J. and Tommelein, I.D. (1999). "Space Scheduling using Flow Analysis" in Tommelein, I.D. (ed.). *Proc. 7th Annual Conf. International Group for Lean Construction*, 26-28 July, Berkeley, CA, 299-311.

Choo, H.J. and Tommelein, I.D. (2000a). "WorkMovePlan: Database for Distributed Planning and Coordination." *Proc. 8th Ann. Conf. Intl. Group for Lean Construction* (IGLC-8), 17-19 July, held in Brighton, UK.

Choo, H.J. and Tommelein, I.D. (2000b). "Interactive Coordination of Distributed Work Plans." *Proc. Construction Congress VI*, ASCE, 20-22 Feb., Orlando, Florida, 11-20.

Choo, H.J. and Tommelein, I.D. (2001). "Requirements and Barriers to Adoption of Last Planner Computer Tools." Ninth Annual Conference of the International Group for Lean Construction (IGLC-9), 6-8 August, Singapore.

Choo, H.J., Tommelein, I.D., Ballard, G., and Zabelle, T.R. (1999). "WorkPlan: Constraint-based Database for Work Package Scheduling." ASCE, *J. of Constr. Engrg. and Mgmt.*, 125 (3) 151-160.

Microsoft (1999). *Synchronization*. Microsoft Corp., Redmond, WA, http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dao360/html/ rpldefsynchronization.asp.

Microsoft (2000a). Microsoft Access, Microsoft Corporation, Redmond, WA, http://www.microsoft.com/office/access/.

Microsoft (2000b). Microsoft Exchange Server, Microsoft Corporation, Redmond, WA, http://www.microsoft.com/exchange/.

Microsoft (2000c). Microsoft Project, Microsoft Corporation, Redmond, WA, http://www.microsoft.com/office/project/

Microsoft (2000d). Microsoft Visio, Microsoft Corporation, Redmond, WA, http://www.microsoft.com/office/visio/.

Russell, A. and Froese, T. (1997). "Challenges and a Vision for Computer-integrated Management Systems for Medium-sized Contractors." *Can. J. Civ. Engrg.*, 24 (2) 180-190.

Scherer, R., Keller, M., and Menzel, K. (2002). "Personalized Cross-Project Workflow Approach For Multi-Project Participants." *3rd Intl. Conf. on Concurrent Engrg. in Construction* (CEC02), 1-3 July, Univ. of California, Berkeley, CA, pp. 41-50, http://www.ce.berkeley.edu/~tommelein/concur02/SchererEtAl.pdf

Tommelein, I.D. and Zouein, P.P. (1993). "Interactive Dynamic Layout Planning." *J. Constr. Engrg. and Mgmt.*, ASCE, 119 (2) 266-287.

VRML (1995). *The Virtual Reality Modeling Language, Version 1.0 Specification, 26-MAY-95*, available at http://www.virtpark.com/ theme/vrml/, visited 7/16/99.

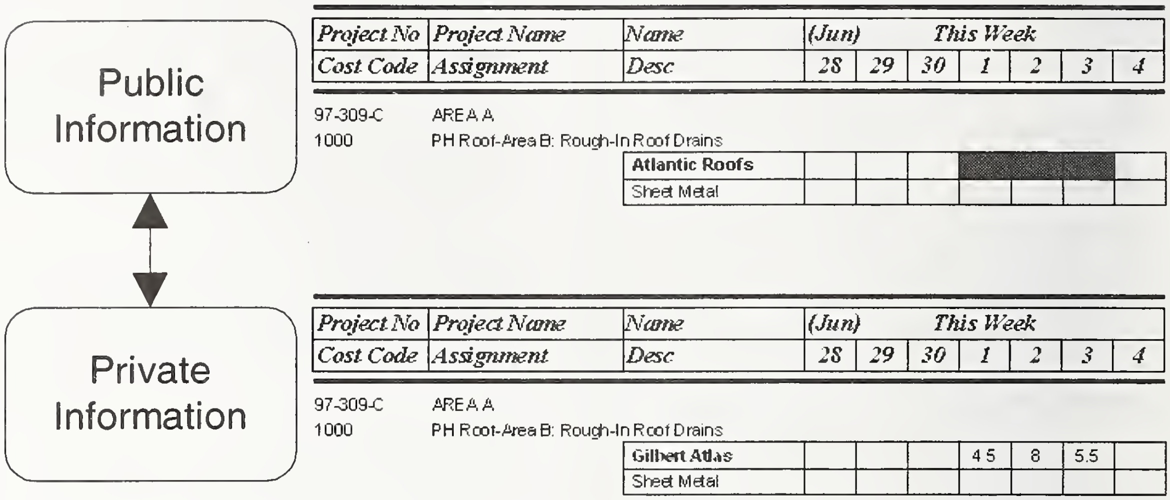


Figure 1. Relationship between Private Information vs. Public Information

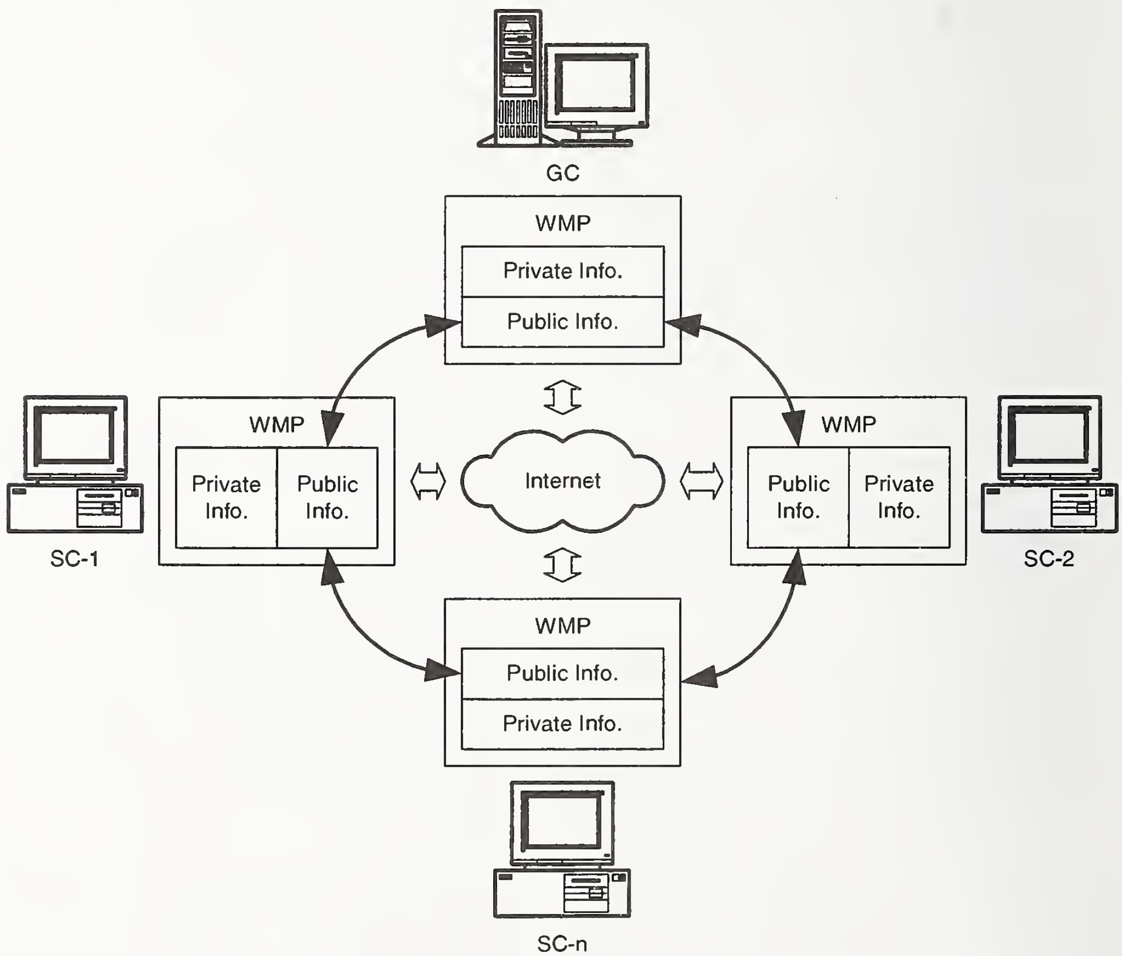


Figure 2. WorkMovePlan (WMP) Synchronization Scheme

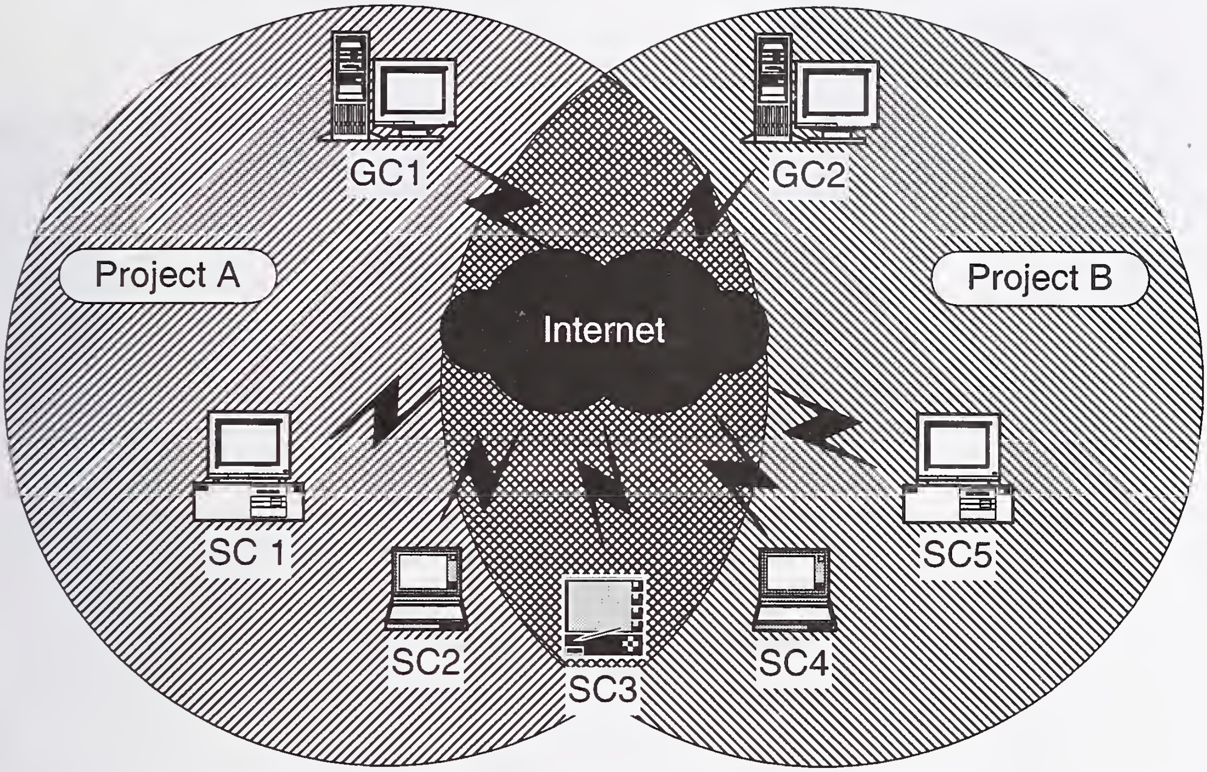


Figure 3. WorkMovePlan's Multi-project Scheduling Scheme

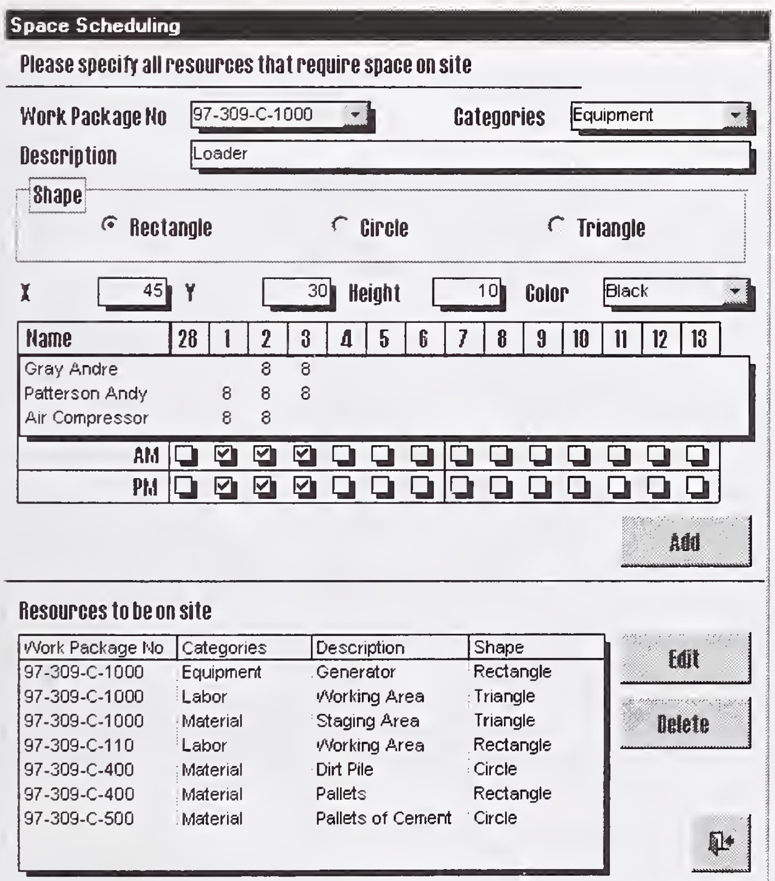


Figure 4. Space Scheduling Screen 1

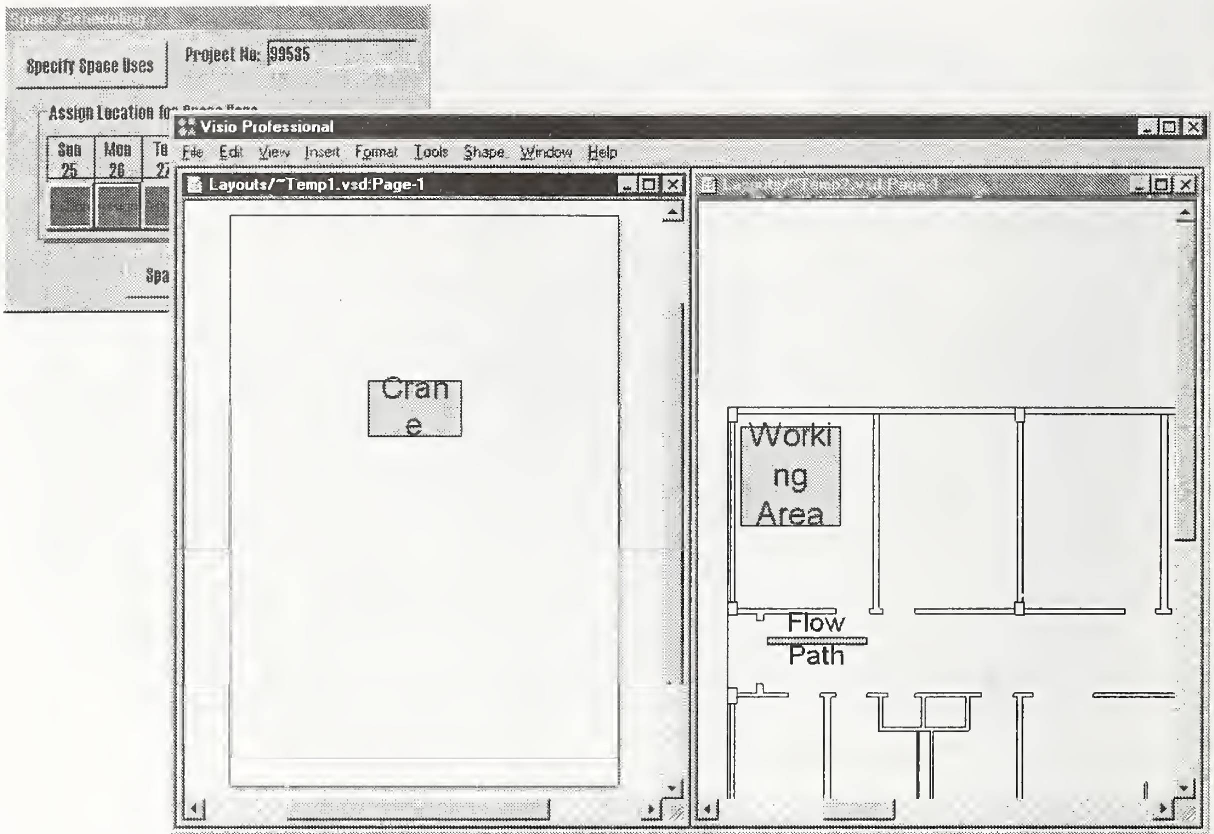


Figure 5. Space Scheduling Screen 2

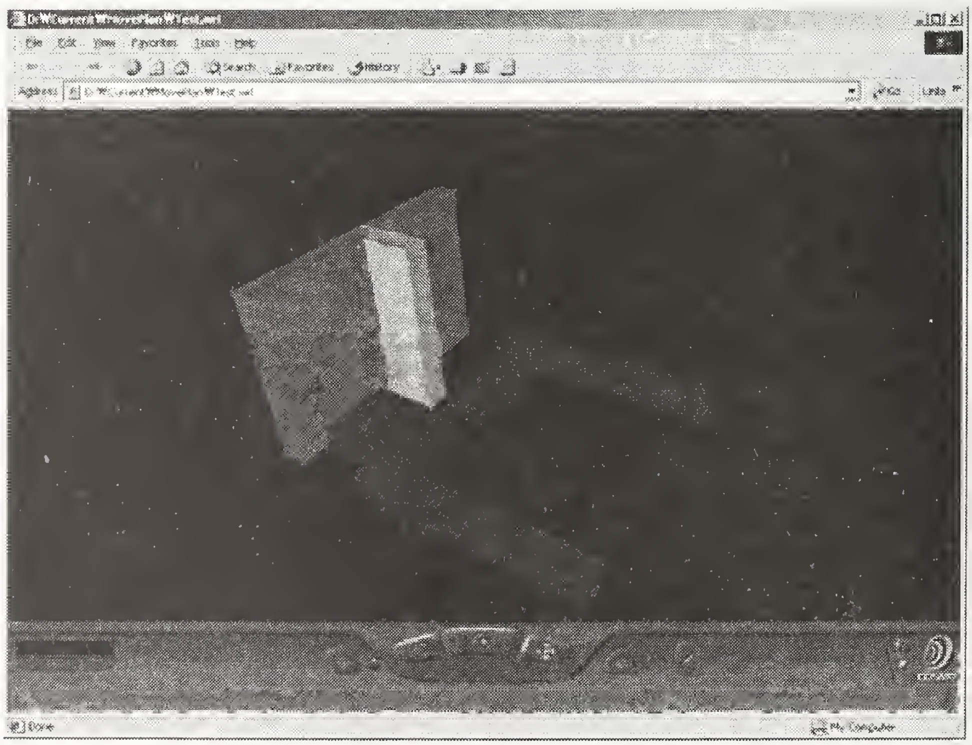


Figure 6. Sample Site Layout using VRML

Project Performance vs. Use of Technologies at the Work Function Level

by

James T. O'Connor[1] and Li-Ren Yang[2]

ABSTRACT: An industry-wide survey was used to collect project data from more than 200 capital facility projects on the issue of technology usage at the work function (WF) level and overall project success. Findings pertaining to associations between project success and technology usage at the work function level are discussed. The project success variables analyzed include project schedule success and project cost success.
Research hypotheses analyzed in this study are presented as follows: 1) High-Tech WFs vs. Project Schedule Success, 2) Low-Tech WFs vs. Project Schedule Failure, 3) High-Tech WFs vs. Project Cost Success, and 4) Low-Tech WFs vs. Project Cost Failure. Project schedule success or failure is particularly leveraged with technology usage or lack thereof for developing scope of work, acquiring & responding to shop drawings, communicating Requests for Information & response, providing feedback about cost & schedule impacts from changes, using as-built information in operator training, and updating as-built drawings. Project cost success or failure is particularly leveraged with technology usage or lack thereof for monitoring facility energy consumption.

KEYWORDS: cost success, schedule success; technology usage; work function; work function characteristics

1. INTRODUCTION

1.1 Study Background, Study Objectives, and Scope Limitations

This paper presents findings pertaining to associations between project success and technology usage at the work function level. The data upon which these statistics are based were collected from more than 200 capital facility projects in the lower 48 states of the U.S. between October 1998 and August 1999.

Technology usage metrics analyzed include those for High- and Low-Tech WFs. High-Tech and Low-Tech WFs are associated with the highest levels of technology utilization (Level 3) and the lowest levels of technology utilization (Level 1) in executing work functions for the subject project, respectively.

The project success variables analyzed include final performance of the projects in terms of schedule and cost success. With respect to the schedule success variable, schedule success is defined to have occurred when the actual project completion date was significantly *earlier* than planned. Schedule failure occurs when the actual project completion date was significantly *later* than planned. For the cost success variable, cost success is defined to have occurred when the total installed cost was significantly *under* authorized budget. Cost failure occurs when the total installed cost was significantly *over* authorized budget.

1.2 Research Hypotheses

Research hypotheses analyzed in this study are detailed as follows: 1) Higher levels of project schedule success are associated with certain WFs when High-Tech approaches are applied to those WFs, 2) Lower levels of project schedule success are associated with certain WFs when Low-Tech approaches are applied to those WFs, 3) Higher levels of

[1] Professor, Department of Civil Engineering, ECJ 5.2, The University of Texas, Austin, TX 78712; PH 512-471-4645; jtoconnor@mail.utexas.edu
[2] Graduate Research Assistant, Department of Civil Engineering, ECJ 5.2, The University of Texas, Austin, TX 78712; lry@mail.utexas.edu

project cost success are associated with certain WFs when High-Tech approaches are applied to those WFs, and 4) Lower levels of project cost success are associated with certain WFs when Low-Tech approaches are applied to those WFs.

1.3 Methodology

Salient aspects of the research methodology are presented as follows:

- Small projects (<$5mil.) were excluded from the analysis.
- High-Tech/High Schedule Success WFs involve significantly more technology usage and are associated with a higher rate of schedule success.
- Low-Tech/Low Schedule Success WFs involve significantly less technology usage and are associated with a lower rate of schedule success.
- High-Tech/High Cost Success WFs involve significantly more technology usage and are associated with a higher rate of cost success.
- Low-Tech/Low Cost Success WFs involve significantly less technology usage and are associated with a lower rate of cost success.
- Project schedule success or failure is particularly leveraged with technology usage or lack thereof for the work functions pertaining to both High-Tech/High Schedule Success WFs and Low-Tech/Low Schedule Success WFs.
- Project cost success or failure is particularly leveraged with technology usage or lack thereof for the work functions pertaining to both High-Tech/High Cost Success WFs and Low-Tech/Low Cost Success WFs.

2. SCHEDULE SUCCESS FINDINGS

2.1 High-Tech/High Schedule Success WFs

Table 1 presents High-Tech WF descriptive statistics according to project schedule performance. High-Tech/High Schedule Success WFs involve significantly more technology usage and are associated with a higher rate of schedule success, so High-Tech/High Schedule Success WFs include the following work functions (presented in order of significance):

- Develop scope of work
- Model user's process
- Conduct needs analysis
- Prepare milestone schedule
- Train facility operators
- Use as-built information in operator training
- Provide feedback about cost and schedule impacts from changes
- Earthwork & grading
- Acquire and respond to shop drawings
- Develop detailed construction schedule
- Communicate Requests for Information & response

2.2 Low-Tech/Low Schedule Success WFs

Table 2 presents Low-Tech WF descriptive statistics according to project schedule performance. Low-Tech/Low Schedule Success WFs involve significantly less technology usage and are associated with a lower rate of schedule success, so Low-Tech/Low Schedule Success WFs include the following work functions (presented in order of significance):

- Provide feedback about cost and schedule impacts from changes
- Link between supplier cost quotes and cost estimate
- Request facility maintenance or modifications
- Update as-built drawings
- Use as-built information in operator training
- Develop scope of work
- Submit contractor's request for payment
- Provide elevated work platform
- Acquire and respond to shop drawings
- Detect physical interferences
- Acquire and record material lab test results

3. COST SUCCESS FINDINGS

3.1 High-Tech/High Cost Success WFs

Table 3 presents High-Tech WF descriptive statistics according to project cost performance. High-Tech/High Cost Success WFs involve significantly more technology usage and are associated with a higher rate of cost success, so High-Tech/High Cost Success WFs include the following work functions (presented in order of significance):

- Monitor facility energy consumption
- Monitor environment impact from operations
- Fabricate roof trusses
- Design HVAC systems
- Design electrical systems
- Monitor equipment operations
- Conduct needs analysis
- Track the inventory of materials on site
- Prepare floor plans
- Link between quantity survey and cost estimate
- Model user's process

3.2 Low-Tech/Low Cost Success WFs

Table 4 presents Low-Tech WF descriptive statistics according to project cost performance. Low-Tech/Low Cost Success WFs involve significantly less technology usage and are associated with a lower rate of cost success, so Low-Tech/Low Cost Success WFs include the following work functions (presented in order of significance):

- Use as-built information in operator training
- Update as-built drawings
- Track design progress
- Document budget assumptions
- Train facility operators
- Providing feedback about cost and schedule impacts from changes
- Update as-built drawings
- Owner payment to contractor
- Monitor facility energy consumption

4. WORK FUNCTION CHARACTERISTICS

Additional analyses of the data are on going and pertain to Work Function Characteristics (WFCs). WFCs are differentiae that characterize the 68 work functions. A total of 31 WFCs based on 6 categories (i.e., WF procedures, Time/Space/Cost, Information & data, Management, WF product, and Human resource) were developed by O'Connor and Won to classify work functions by their attributes. WFC analysis reveals characteristics common to a specific work function group. This approach helps explain why different levels of technology usage exist and why specific technologies and tools are in more demand. Details associated with WFCs and discussion regarding these analyses are included in Won's dissertation.

5. CONCLUSIONS

5.1 Analysis Results for Schedule Success

Presented in order of significance, High-Tech/High Schedule Success WFs include the following work functions:

- Develop scope of work
- Model user's process
- Conduct needs analysis
- Prepare milestone schedule
- Train facility operators
- Use as-built information in operator training

Presented in order of significance, Low-Tech/Low Schedule Success WFs include the following work functions:

- Provide feedback about cost and schedule impacts from changes
- Link between supplier cost quotes and cost estimate
- Request facility maintenance or modifications
- Update as-built drawings
- Use as-built information in operator training
- Develop scope of work

Attention should be paid to the work functions pertaining to both High-Tech/High Schedule Success WFs and Low-Tech/Low Schedule Success WFs. Project schedule success or failure is particularly leveraged with technology usage or lack thereof for these work functions:

- Develop scope of work
- Acquire and respond to shop drawings
- Provide feedback about cost & schedule impacts from changes
- Use as-built information in operator training

5.2 Analysis Results for Cost Success

Presented in order of significance, High-Tech/High Cost Success WFs include the following work functions:

- Monitor facility energy consumption
- Monitor environment impact from operations
- Fabricate roof trusses
- Design HVAC systems
- Design electrical systems
- Monitor equipment operations
- Conduct needs analysis
- Track the inventory of materials on site

Presented in order of significance, Low-Tech/Low Cost Success WFs include the following work functions:

- Use as-built information in operator training
- Update as-built drawings
- Track design progress
- Document budget assumptions

Project cost success or failure is particularly leveraged with technology usage or lack thereof for the work function "Monitor facility energy consumption."

6. RECOMMENDATIONS FOR FUTURE RESEARCH

Recommendations for future study are offered:

- Work function characteristics associated with High-Tech/High Schedule Success WFs, Low-Tech/Low Schedule Success WFs, High-Tech/High Cost Success WFs, and Low-Tech/Low Cost Success WFs may help further explain project success.
- Any future similar survey should involve expansion of assessment levels from 3 to 4 in order to improve resolution of estimates of technology usage.

7. REFERENCES

[1] O'Connor, J. T., Kumashiro, M. E., Welch, K. A., Hadeed, S. P., Braden, K. E., and Deogaonkar, M. J. (2000). "Project- and phase-level technology use metrics for capital facility projects." *Report No.16*, Center for Construction Industry Studies, University of Texas, Austin, Texas.

[2] O'Connor, J. T., and Won, S. (2001). "Work function-level technology use metrics for capital facility projects." *Report No.18*, Center for Construction Industry Studies, University of Texas, Austin, Texas.

[3] O'Connor, J. T., and Won, S. (2001). "Understanding Demand for Technology via Work Function Characteristics", *Proc., Special Conference on Fully Integrated and Automated Project Processes*, Virginia Polytechnic and State University, Blacksburg, Virginia, 13-24.

[4] O'Connor, J. T., and Yang, L. (2002). "Capital facility project success and technology usage." *Report No.22*, Center for Construction Industry Studies, University of Texas, Austin, Texas.

[5] Won, S. (2002). "A model for work function-based prioritization of technologies for capital projects." PhD thesis, University of Texas, Austin, Texas.

Table 1. High-Tech WF Descriptive Statistics by Project Schedule Performance

ID	WF	% of Responses at Level 3		Δ %	Rank
		% of Projects with Schedule Success	% of Projects with Schedule Failure		
1.01	Conduct needs analysis	33	0	33	3
1.02	Develop scope of work	36	0	36	1
1.03	Model user's process	42	8	34	2
1.05	Prepare milestone schedule	40	10	30	4
3.09	Acquire & respond to shop drawings	22	0	22	9
4.01	Develop detailed construction schedule	33	11	22	9
4.09	Communicate Requests for Information & response	25	5	20	11
4.10	Cost & schedule impacts from changes	24	0	24	7
5.02	Earthwork & grading	24	0	24	7
6.02	Train facility operators	33	8	25	5
6.03	Use as-built information in operator training	33	8	25	5

Table 2. Low-Tech WF Descriptive Statistics by Project Schedule Performance

ID	WF	% of Responses at Level 1		Δ %	Rank
		% of Projects with Schedule Failure	% of Projects with Schedule Success		
1.02	Develop scope of work	26	0	26	6
2.11	Detect physical interferences	41	20	21	10
3.04	Link between supplier cost quotes and cost estimate	50	19	31	2
3.09	Acquire & respond to shop drawings	56	33	23	8
4.10	Cost & schedule impacts from changes	78	29	49	1
4.14	Submit contractor's request for payment	53	29	24	7
5.06	Provide elevated work platform	69	46	23	8
5.09	Acquire & record material lab test results	50	29	21	10
6.03	Use as-built information in operator training	62	33	29	5
6.07	Request facility maintenance or modifications	50	20	30	3
6.08	Update as-built drawings	43	13	30	3

Table 3. High-Tech WF Descriptive Statistics by Project Cost Performance

ID	WF	% of Responses at Level 3		Δ %	Rank
		% of Projects with Cost Success	% of Projects with Cost Failure		
1.01	Conduct needs analysis	25	0	25	6
1.03	Model user's process	21	0	21	11
2.05	Prepare floor plans	44	20	24	9
2.08	Design electrical systems	47	18	29	5
2.09	Design HVAC systems	47	14	33	3
3.03	Link between quantity survey and cost estimate	31	8	23	10
4.06	Track the inventory of materials on site	25	0	25	6
5.07	Fabricate roof trusses	33	0	33	3
6.06	Monitor equipment operations	38	13	25	6
6.09	Monitor facility energy consumption	70	14	56	1
6.10	Monitor environment impact from operations	43	0	43	2

Table 4. Low-Tech WF Descriptive Statistics by Project Cost Performance

ID	WF	% of Responses at Level 1		Δ %	Rank
		% of Projects with Cost Failure	% of Projects with Cost Success		
2.10	Document budget assumptions	36	12	24	4
2.14	Track design progress	42	17	25	3
4.10	Cost & schedule impacts from changes	75	55	20	6
4.13	Update as-built drawings	50	30	20	6
4.15	Owner payment to contractor	63	43	20	6
6.02	Train facility operators	63	40	23	5
6.03	Use as-built information in operator training	67	40	27	1
6.08	Update as-built drawings	40	14	26	2
6.09	Monitor facility energy consumption	29	10	19	9

DEVELOPMENT OF AN INFORMATION MODEL TO ENHANCE INTEGRATION AND COORDINATION IN THE CONSTRUCTION PROJECTS

Amaury A. Caballero Ph.D., PE.; Syed M. Ahmed Ph.D.; Salman Azhar; and Maribel Barcala

Florida International University
10555 W. Flagler Street, Miami, Florida 33174, USA

ABSTRACT: Construction is generally believed to be a fragmented industry and increased integration and coordination among different processes and parties is considered by many experts as one of the ways that can resolve most of the problems created by fragmentation. Process inefficiency and a lack of effective quality control in the products of construction are often attributed to fragmentation. The recent developments in information and communication technology (ICT) have provided an opportunity to increase the level of integration among different construction processes by improving the flow of information. This paper examines the possibilities of increasing integration in the construction industry by integrating the planning, design, and construction processes. In addition, an information model is proposed to integrate the different phases in the construction operations and to increase interaction among the stakeholders.

KEY WORDS: Construction, Information Technology, Coordination, Integration, Modeling

1. INTRODUCTION

The construction industry is highly fragmented as compared to other industries. This may have caused significant low productivity, cost and time overruns, conflicts and disputes, resulting in claims and time consuming litigation. (Latham 1994).

The fragmentation problem is further compounded by the fact that the construction process typically involves several disciplines, e.g. architects, structural engineers, building services (HVAC) engineers, quantity surveyors, contractors, sub-contractors, material suppliers etc., collaborating for relatively short periods in the design and construction of a facility. Until fairly recently, these disciplines tended to work independently, while making decisions that affect the others (Anumba, 2000).

Another facet of the fragmentation problem is the fact that the construction projects whether they be buildings, bridges, dams, or offshore structures usually involve many stages, starting from the establishment of the client's requirements through to design, construction, utilization and eventual disposal of the facility. These stages of the project's life cycle and the associated activities and tasks are often undertaken as discrete processes, with only limited integration of data/information, participants, tools and procedures etc (Anumba, 2000).

Some of the consequences of the fragmentation problem include (Amor and Anumba, 1999):

- Inadequate capture, structuring, prioritization and implementation of client needs.
- The fragmentation of design, fabrication and construction data, with data generated at one project not being readily re-used downstream.
- Lack of integration, coordination and collaboration between the various functional disciplines involved in the life-cycle aspects of the project.
- Lack of true life cycle analysis of projects (including costing, maintenance etc.).
- Poor communication and design intent and rationale, which leads to

unwarranted design changes, unnecessary liability claims, increase in design time and cost, and inadequate pre- and post-design specifications.

To achieve the benefits obtained from a good information system, it will be necessary to clarify:

- What information is needed, in which form and during which project stage?
- What decisions will be made and when?
- Who should contribute to these decisions?
- Who should communicate to whom?

The purpose of this research was to find the answers to these questions and present them in the form of a graphical model, which could integrate the different construction stages. A number of research projects have addressed aspects of integration among two or more stages in the building project life cycle. However, there are presently no integration models available that are applicable to the whole project life cycle from conception to demolition. Nevertheless the feasibility and desirability of fully integrated lifecycle models has been restricted to small models. There is a gap in research particularly into the development of appropriate integrated models for the planning, conceptual design, and demolition stages of the project lifecycle. An important challenge in lifecycle integration is ensuring that the models are consistent, persistent, and able to convey data as well as the underlying design intent and rationale. (Amor, 1999; Anumba, 2000).

Important opportunities for improving performance of engineering and construction projects may be obtained through the integration of planning, design, and construction phases. With the construction integration process, suppliers have a growing influence on the design, which results in designs that better fit the construction needs. This process has been defined as the continuous and interdisciplinary sharing of goals, knowledge, and information among all project participants (Gilbertus, 1997).

The recent efforts of rationalizing the industry reached the point to integrate planning, design, fabricating, and assembly process like manufacturing industry. The previously separated design information and construction process planning are combined and integrated as a big construction system. That conceptual progress is considered (Hasegawa, 2000) as the trigger of construction automation.

2. OBJECTIVES

The main objective of this paper is to present a model that integrates the project through the construction stages in a building. This is not a feasibility study nor is it a software architecture. The authors propose a framework for the flow of information between the different construction process participants. It is basically a logic diagram (Figures 1 and 2) that can be used to develop an architecture. These stages are clearly defined in previous works. For any building it is possible to identify four construction stages and one operation stage. The construction stages are defined as: (Eastman, 1993):

1. Feasibility study
2. Design
3. Construction planning
4. Construction

3. METHODOLOGY

The model has been improved during two years with the help of the Construction Management Program graduate students working in the construction industry. All of our graduate students are working full-time in the local construction industry in responsible managerial positions and have an average of over 10 years of experience in the construction industry. The authors believe their input is therefore, very relevant to real construction projects. The goal was to receive feedback from them that could help us to refine it. The following guidelines were included to help in the graduate students work.

1. All the parties will work on the product from the early stages up to the production stage like a team. The previously separated design information and construction process planning stages

have to be combined and integrated as a fully integrated construction system.

2. The process will be changed interactively and according to the needs.

3. All the information should be in a standard format recognized by all the professionals.

The data flow diagram for each stage in a construction project is shown in Figure 1, and the interaction between the different parties is shown in figure 2. The data model exhaustively defines the construction project through its entire construction cycle. The data includes graphic representation in four dimensions, technical and cost calculations, scheduling, etc. The ultimate goal of this model is to form a software bridge between mathematical coding - the language that computers read - and the characters and pictures that human can see and understand.

4. MODEL DISCUSSION

The first stage in any construction project is the feasibility study. The feasibility study is the generator of the building model and thus influences the design and later stages. This stage also plans and sets goals at a general level for all the other stages. It defines the purposes of the building project and assesses if the resources are appropriately matched with the project scope. At this stage, the costs are balanced with the function of the building. The planning at this level of the building model often involves developing many different feasibility models and comparing them in different dimensions.

The design stage involves the translation of functional criteria developed in the feasibility stage into detailed descriptions of the building project, including the preparation of detailed drawings and specifications. The advantage of this model is that since all the parties are involved in this process, different design alternatives could be easily evaluated and the bets and optimized design can be selected,

In the construction planning stage the bidding, the impact on cost/schedule, and the construction plan are completed. The construction stage executes the construction plan. During this stage, if there is any change in the design or specifications, it could be easily incorporated by the consultation of the owner and the designer.

Due to the cheap availability of the information technology, this model could be easily converted into a software system, which can store and manage all the information. Such software could be coupled with the state-of-the-art web-based technologies. In this way, all the information will be available on-line and accessible to all the project parties regardless of time and space limitations.

5. CONCLUSIONS

A model to increase interaction among stakeholders in the construction process has been presented. The proposed model attempts to deal with all the involved parties from the onset of the project in a manner similar to a construction team. This is a great advantage for the project itself because all changes can be discussed among the different parts of the team before these have to be implemented. This is also very helpful to the management in making better decisions about any upgrade of the project. The model may also contribute to reducing the level of risks during the construction stage.

One effect of the construction integration process is that the suppliers will have more influence on the design process, which results in designs that better fit the construction needs. The recent efforts of rationalizing the industry has reached the stage where it is necessary to integrate planning, design, fabricating, and assembly process like the manufacturing industry. Managerial functions like engineering, management control, contract administration, and others, move towards the integration.

One limitation of this study could be that the feedback was generated only from graduate students. It is recommended that future studies could be expanded to include more diverse participants from the construction industry.

6. REFERENCES

Amor R. & C. J. Anumba (1999, August). *A Survey and Analysis of Integrated Project Databases.* Concurrent Engineering Construction: Challenges for the New Millennium, CIB Publication 236, Espoo, Finland. 217-227.

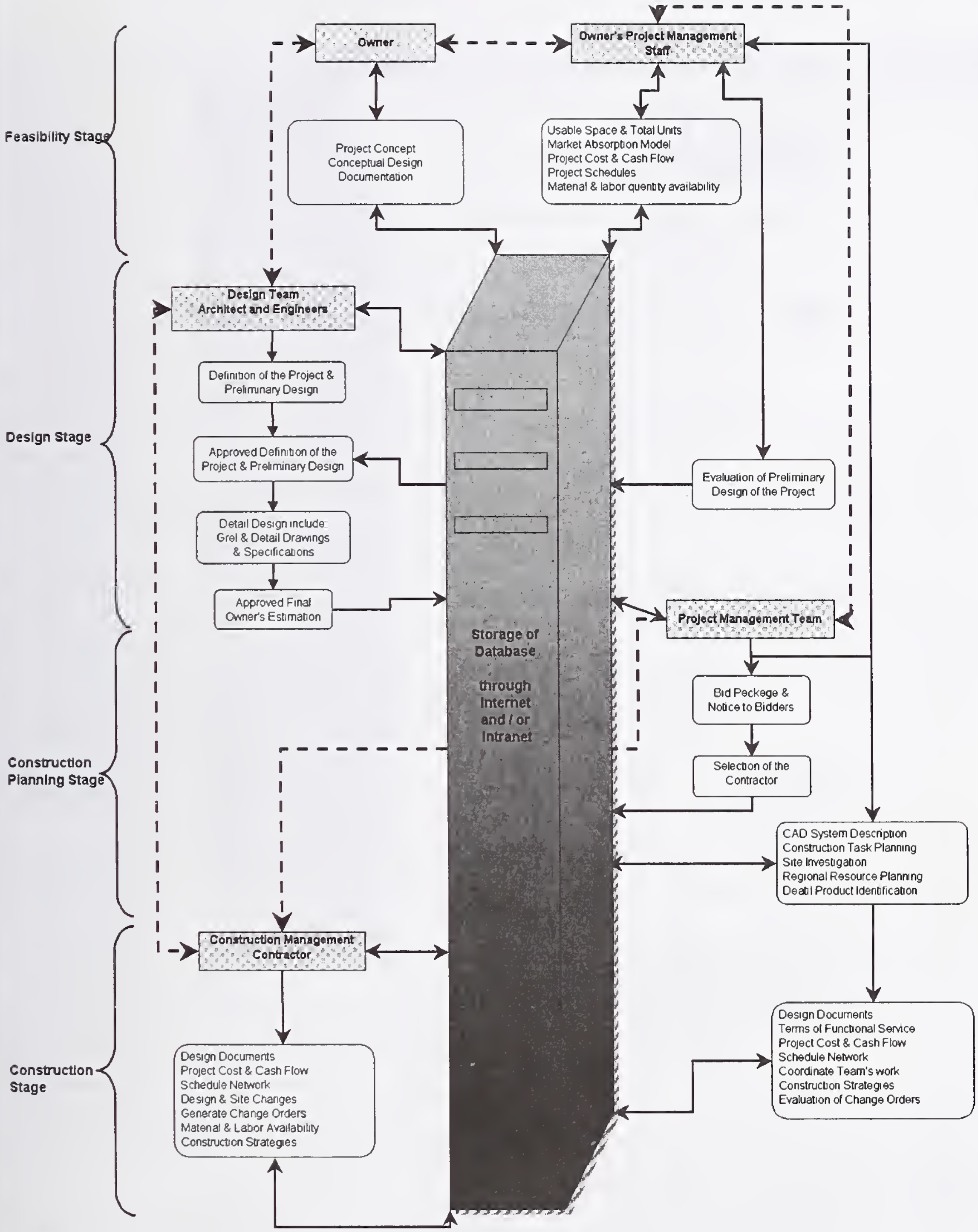
Anumba C. J. (2000). *Integrated Systems for Construction: Challenges for the Millennium.* Proceedings of the International Conference on Construction Information Technology (INCITE) 2000, Hong Kong, 78-92.

Eastman, C. M. (1993, August). *Life Cycle Requirements for building product models.* Proceedings of the 1st International Conference on Management of Information Technology for Construction. Singapore.

Gilbertus T. (1997, June). *Automating Communication in Civil Engineering*. Journal of Construction Engineering and Management. V-123, No. 2. 113-120.

Hasegawa Y. A. (2000). *New Wave of Construction Automation and Robotics in Japan.* Proceedings of the 17th International Symposium on Automation and Robotics in Construction ISARC 2000. Taipei, Taiwan, k15 – k20.

Latham, M. (1994) *Constructing the Team.* London: Her Majesty Stationary Office (HMSO).

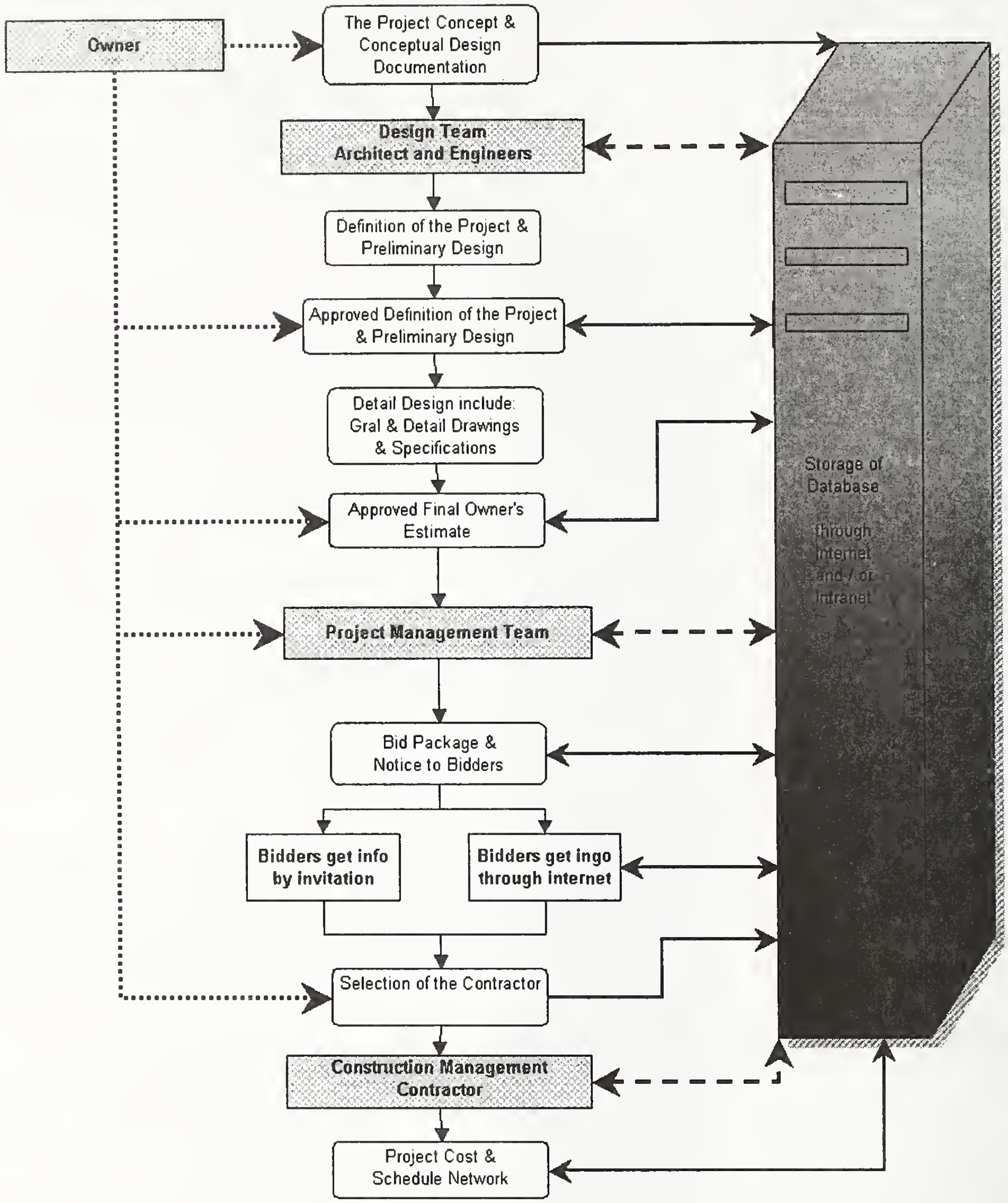


NOTES:

The dashed line represents the interaction between the parties.

The continuous line shows how the data is sent to or received from the database storage.
However, the interconnection between the processes take place in the database.

Figure 1. Data Flow Diagram for each Stage in a Construction Project



NOTE:
The direction of the arrows indicate how the information is flowing through the system.
⋯⋯▶ The doted line represents the owner's actions through the server.
------▶ The dashed line indicates the interactions of the parties through the database.
———▶ The continuous line shows how the data is sent to or received from the database storage.
We can visualize the proposed management process by using a database storage on Internet and / or Intranet.
The information is available to any party at anytime. Therefore, it would benefit communication between the parties during the bid process.
Bidders can get information easily reglardless of where they are located.
In case of design changes, the Project Management Team can access information instanstaneously.
Because all parties involved have access to the same database, they have the same information source.

Figure 2. Interaction Between Parties

Development of an Integrated Cost Estimation and Cost Control System for Construction Projects

by

Salman Azhar, Syed M. Ahmed and Amaury A. Caballero

Florida International University
10555 W. Flagler Street, Miami, Florida 33174, USA

ABSTRACT

Today, most of the construction companies are using computer-aided softwares for cost estimation and cost control. The cost estimation is done before the construction begins while the cost control is performed during the construction period. The organization which controls the cost could be different from the organization which prepares the cost estimate. Moreover, different softwares are available for cost estimation and cost control. This may complicate the cost control process due to inconsistencies between the cost estimation and cost control data. The purpose of this study is to develop a system which could integrate the cost estimation and cost control processes. Such integration enables the transfer of cost estimation data automatically to the cost control process. Therefore, complete information regarding the cost status of the project can be accessed at any time. Systems Analysis and Design (SAD) approach using relational database management system is used to develop the integration model. MS Access™ is used as a physical tool to implement and test the system. This paper provides a glimpse of the requirements analysis, system development and implementation stages.

KEYWORDS

Cost Estimation; Cost Control; Cost Integration; Cost Engineering; Construction Softwares, Data Modeling

1. INTRODUCTION

Cost engineering consists of cost estimation and cost control. Estimation is a methodology for forecasting and predicting cost and expenditures of a future project and to produce a budget. Cost control is a process to hold expenditures within the budget by monitoring and appraisal of the cost performance (Clark and Lorenzoni, 1997).

In most construction projects, cost estimation and cost control are performed by different people (e.g. estimation is done by quantity surveyors or estimators while cost control is the responsibility of the owner or contractor's cost engineers) and sometimes by different organizations using different techniques and/or softwares (Kibler, 1992). Due to these reasons, the cost estimation and cost control data are usually available in different forms and at different levels of detail. This inconsistency of data results in lack of accuracy and untimely generation of cost control reports for any corrective decision making (Cagle, 1991).

However, if the estimation and control processes are integrated, there will be a communication of information between the two processes. The results of the cost control process can be used as a feedback to the estimation process for future projects, while a good estimate may make the control process more effective.

Therefore, this study is conducted with the objective of developing an integrated cost estimation and control system which could facilitate the transfer of information between the two processes for current and future projects. The system identifies the linkages (common attributes) between the two processes and presents a data flow schema to implement it in the real life using any database management software such as MS Access™ or Oracle™.

2. METHODOLOGY

Due to its inherent simplicity, systems analysis and design (SAD) approach is selected to develop the integration model and the prototype software (Whitten et al., 2001). The framework selected for this research is outlined in Figure 1, which consists of three steps with interlinked inputs and outputs.

2.1 Requirements Analysis

There are three types of requirements which were identified for developing the cost integration model. These are cost estimation software requirements, cost control software requirements and the requirements for their integration. For this purpose, the existing cost estimation and control softwares were explored to identify their vital attributes which could be used for integration. Such attributes included the cost and work activities coding system, techniques to measure different processes, and input and output formats.

2.2 Integrated Model Development

The integrated model development begins by developing the individual data schemas for cost estimation and cost control parts. This is done by plotting the logical dataflow diagrams (DFD) for each part, which provided a key to integrate both parts. This process is briefly illustrated in Figure 2.

In the next step, the data requirements of the integrated cost estimation and control processes are identified. These data requirements are then combined to prepare an integrated data model (i.e. an Entity-Relation (E-R) diagram) which serves as a basis to construct prototype software.

2.3 System Development and Testing

MS Access™ is used as commercial-of-the-shelf tool to develop prototype software using the integrated model developed in the second step. After the prototype is created, its testing is conducted to determine any errors, problems and feedback for improvement.

3. REQUIREMENTS ANALYSIS

3.1 Coding System

CSI MASTERFORMAT is used for cost estimation and control coding which assigns a specific activity id to each individual activity (O'Connor, and Caraway, 1993). The activity id is divided into 4 parts: 1_id, 2_id, 3_id, and item_id representing major item, minor item, type of construction and the location. For example, activity id 03 010 020 001 shows formwork and support for concrete columns in the first floor of the building.

3.2 Cost Estimation and Control Processes

Cost estimation includes estimates for materials, labor, equipment and subcontracts. For the sake of simplicity, lump sum contract is used in this study. Therefore, the format of estimating model consists of work package cost plus markup and overhead as a percentage of total work packages cost for overall project.

The cost control model is based on the earned value concept, consisting of three elements: actual cost of work performed (ACWP), budgeted cost of work performed (BCWP) and budgeted cost of work scheduled (BCWS).

ACWP for material = total amount paid for materials in a work package
ACWP for labor = total amount paid to labor for a work package
ACWP for equipment = total amount paid for equipment assigned to a work package

BCWP for material, labor or equipment = percent work progress multiplied by total cost of labor or equipment or material in a work package

BCWS for material, labor or equipment = total estimated cost for labor, equipment, or material at the control time.
On the basis of these parameters, cost variances and schedule variances are calculated to

determine work productivity which is used to measure the rate of work progress.

3.3 Input and Output Formats

The inputs for the system are the quantities of work packages, unit prices, work schedule and the daily progress data. The output will be cost estimates and cost control reports for overall or any specific part of the project.

4. MODEL DEVLOPMENT

4.1 Dataflow Diagrams

One basic requirement in the integrated cost estimation and cost control model is that the estimation data should be automatically transferred to the control part. Keeping this condition in mind, the logical data flow diagrams for both parts are designed. The logical dataflow diagram depicts how the system behaves without any physical implementation and provides a great tool in developing an information system.

Figure 3 illustrates the dataflow diagram (DFD) for estimation part. The 3-D rectangular boxes show the external agents which interacts with the system and provide necessary inputs. These inputs are then passed to different processes labeled in the logical order. Each process is linked with a certain database to store and/or retrieve the information. The most important process in the estimation DFD is process no. 3 which prepares the cost estimation sheet for material, labor, equipment and subcontracts. This estimation sheet which is developed before the construction stage of the project life cycle is used during the construction stage as a reference of cost. Therefore, the estimation sheet is combined with the project scheduling in order to track down the costs on daily basis.

Similarly, figure 4 shows the DFD for the cost control part. The major input is coming from the process no. 3 (i.e. estimation sheet) which is linked to the process no. 7 to map estimation part with the project schedule. The labor time card is used to assess the actual work of an employee for a work package. The labor time card combined with the labor unit price from the estimate storage (storage no.8.0) will produce actual labor cost of work performed (ALCWP), i.e. storage no.9.0.

The material card is used to assess the amount of material used for each work package on site. The material card combined with the material unit price from the estimate storage (storage no. 8.0) will produce an actual material cost of work performed (AMCWP), storage no. 10.0.

Actual equipment cost is similarly calculated from the equipment time card data which is collected from the site combined with the equipment unit price from the estimate storage (storage no. 8.0). However, the equipment idle cost should also be considered by calculating the type of equipment which is not used on that day (storage no.23). The cost of idle equipment will not be used directly in the work package, but it will be added in the calculations of cost variance for total project cost.

The ACWP of material, labor, and equipment will be used to find respective cost variances for each work package performed. This task can be achieved by combining the ACWP for material, labor, and equipment with the budgeted cost of work performed (storage no. 12, 13, and 14). The budgeted cost of work performed can be calculated by combining the quantity of work done and the unit price of material, labor and equipment from the estimate storage (process no. 11).

Finally, the cost variance for complete project can be found by combining the equipment idle cost (storage no.23) with the cumulative of cost variance for material, labor and equipment for the total work package done (storage no. 15, 16, and 17). The cost variance for total project is represented as process no. 18 in Figure 4.

For the purpose of using the cost control data for future projects, a database of labor and equipment productivity is developed. The labor productivity (process no.20) can be assessed from the actual labor cost of work performed storage (storage no. 9) multiplied with the quantity done, and labor unit price from the estimate storage (storage no. 8.0). The productivity of equipment (process no. 21) can be assessed in the same way as labor.

The material waste is also considered for the future project estimation. The material waste could help the future estimator to estimate the waste of materials for a certain work package more accurately. The material waste (process no. 22) can be assessed from the actual material cost of work performed (storage no. 10) combined with quantity take off, material unit price, and conversion of material unit from the estimate data stored in estimate storage (storage no. 8.0)

4.2 Data Model

Based on the data needs as ascertained in the DFD's, a data model is prepared using the relation database concept as shown in Figure 5. The data model is normalized to the 3rd degree to increase the processing speed of the prototype software.

5. SYSTEM DEVELOPMENT AND TESTING

Among the different commercial off-the-shelf softwares, MS Access™ is selected due to its user friendliness, cheap cost and easy availability. As a first step in prototype software development, data *Tables* are designed for material, labor, equipment, subcontracts and work progress. This was followed by *Query* design to establish the relations between different entities. Each process (as shown in Figures 3-4) is interpreted in the form of *Macros* and finally *Forms* are designed to get input and generate reports.

After development, the prototype software is tested on a number of projects to identify any errors and deficiencies (the software results are compared with the manual calculations). This process is repeated until no significant problems are found.

6. CONCLUSIONS

The purpose of integrating the cost estimation and cost control is achieved by using the relational database management system in Microsoft Access. Such integration enables the transfer of cost estimation data automatically to the cost control process. Therefore, complete information regarding the cost status of the project can be accessed at any time. The system provides different tools for controlling the project cost such as productivity data of equipment & labor, and materials waste data. Such information can enable the management to take preventive actions to control the total project costs. The productivity and materials waste data can also be used to prepare estimates for future projects.

7. RECOMMENDATIONS

The integrated cost estimation and control system in this study is designed for lump sum contract projects. There is a need to extend it for different contract types and enhance its usability for multitasks environment.

8. ACKNOWLEDGEMENT

The fundamental concept of this study is taken from the class lecture notes (Information Technology in Construction) of Dr. Chotchai Charoenngam, Associate Professor, Asian Institute of Technology, Bangkok, Thailand.

9. REFERENCES

Cagle, R. F., 1991. Implementing an Owner's Cost Control System, *AACE Transactions,* O.7.

Kibler, B. E., 1992. Integrated Cost and Schedule Control: a piece of cake, *Cost Engineering,* Vol. 34, No. 7

Lorenzoni, D.C., and Clark, F.D., 1997. *Applied Cost Engineering*, 3rd edition, Marcel Dekker, Inc., USA.

O'Connor, J.T., Caraway, P.H., 1993. Need for Specification Format That Accommodated Engineered Projects, *Journal of Construction Engineering and Management*, Vol.119, No.4

Whitten, L.J.; Bentley, L.D.; and Dittman, K.C., 2001. *Systems Analysis and Design Methods*, McGraw-Hill Higher Education, USA.

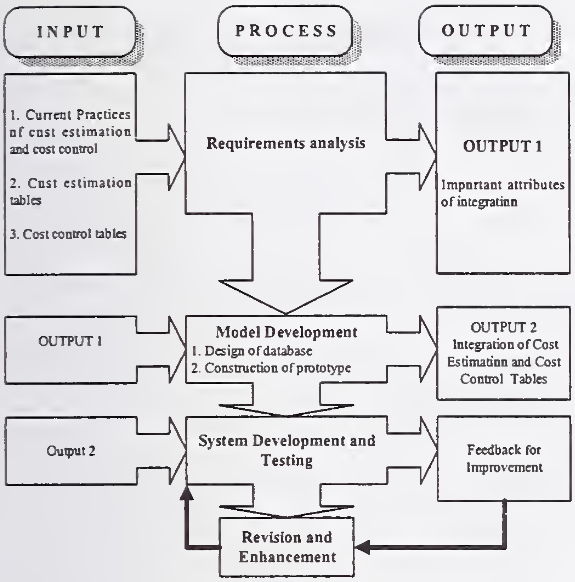


Figure 1. Basic framework to develop an integrated cost estimation and control system

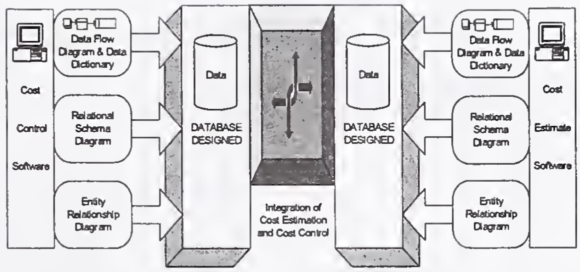


Figure 2. Integrated cost estimation and cost control model development concept

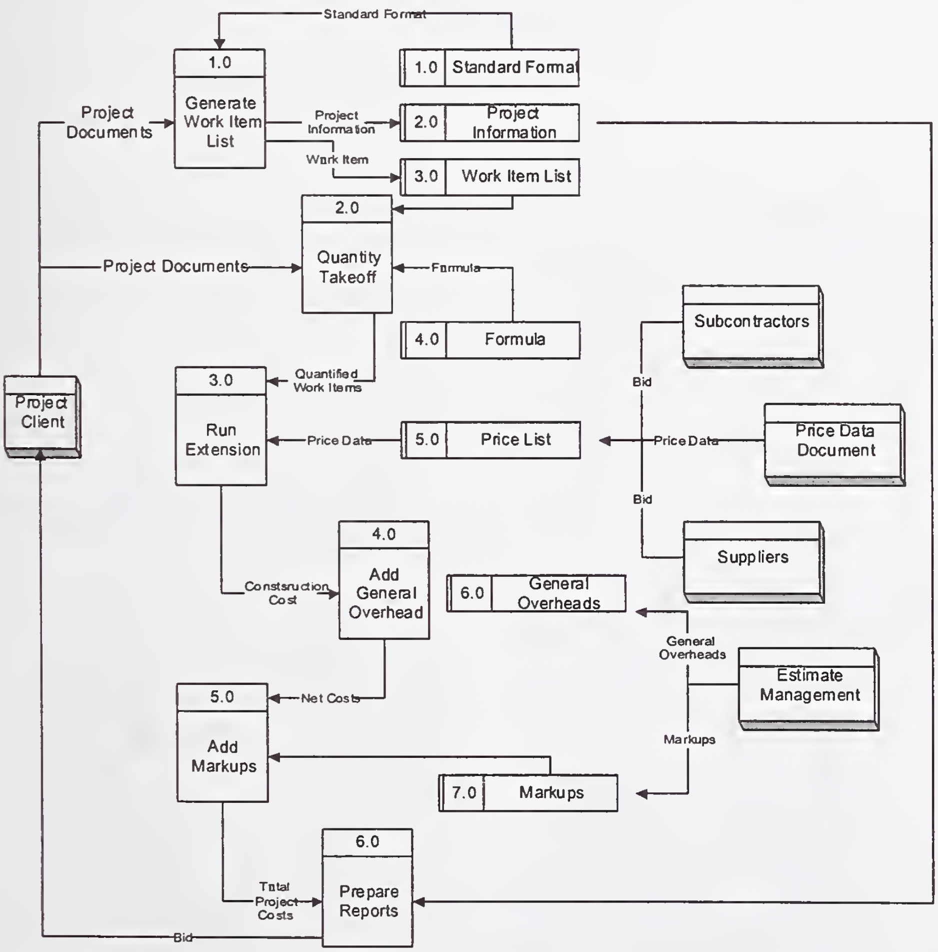


Figure 3. Data flow diagram for cost estimation

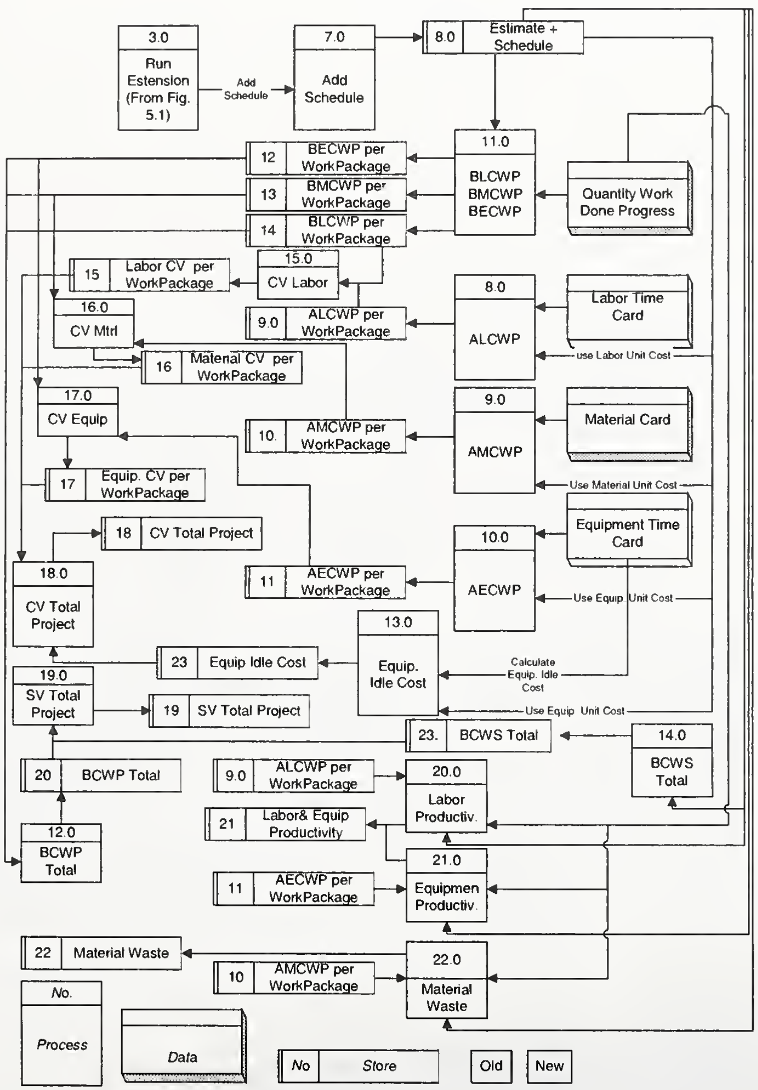


Figure 4. Dataflow diagram for cost control

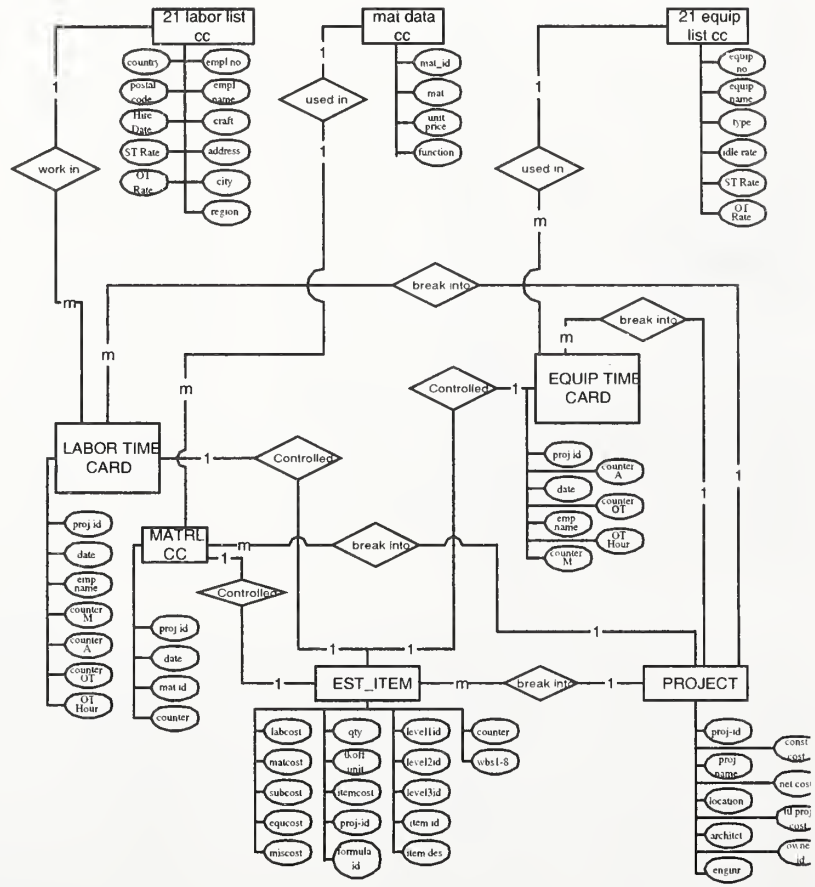


Figure 5. Data model for integrated cost estimation and cost contro system

Automated DSS for Lighting Design of Nighttime Operations in Highway Construction Projects

by

Khaled El-Rayes[1] and Khalied Hyari[2]

[1] Assistant Professor, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 Telephone: (217) 265 0557, Email: elrayes@uiuc.edu, http://cee.ce.uiuc.edu/elrayes

[2] Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 Telephone: (217) 265 5215, Email: hyari@.uiuc.edu

ABSTRACT: Nighttime construction lighting is reported to be a crucial factor affecting quality, safety, cost and productivity of nighttime construction projects. This paper presents an automated decision support system (DSS) for lighting design in highway nighttime construction projects. The DSS is designed to optimize four major objectives: (1) maximizing average illuminance; (2) maximizing lighting uniformity in the work zone; (3) minimizing glare produced by the lighting system; and (4) minimizing cost of the lighting system. Seven decision variables are optimized in this lighting design problem namely: (1) lighting equipment selection; (2) type of lamps; (3) lamp lumen output; (4) mounting height; (5) lighting towers positioning; (6) luminaires aiming angle; and (7) lighting towers rotation. A multi objective evolutionary algorithm (NSGA-II) is used to optimize this lighting design problem. The proposed DSS provides a practical tool for the design and optimization of lighting arrangements in nighttime construction operations.

KEYWORDS: highway construction; lighting; nighttime construction; optimization.

1. INTRODUCTION

An increasing number of highway construction and repair projects throughout the United States are being performed during the off-peak nighttime hours to alleviate construction-related traffic congestions. The use of nighttime operations in highways construction and repair projects is reported to provide many advantages including: (1) reduced impact on the traveling public through reducing congestions and motorist delay; (2) decreased project duration (Hancher and Taylor, 2001); (3) minimized adverse economic impacts of traffic congestion on local commerce particularly for shipping and delivery services; (4) decreased pollution from idling vehicles stopped at construction sites (McCall, 1999); (5) improved work-zone conditions as the smaller amount of traffic at night creates an opportunity to enlarge work zones allowing the concurrent performance of multiple functions; (6) longer working hours at night; (7) enhanced work conditions due to lower temperatures (Shepard and Cottrell, 1985); (8) faster delivery of material to and from the work zone since traffic conditions are better at night, leading to less idle time for both labor and equipment (Price, 1986); and (9) reduced equipment costs (Hancher and Taylor, 2001).

Despite the above advantages, nighttime construction suffers from a number of disadvantages including: (1) decreased visibility for both workers and motorist, causing decreased levels of safety and quality (Shepard and Cottrell, 1985; Hancher and Taylor, 2001); (2) problems in implementing quality control procedures, and decreased quality of workmanship; (3) increased number of drivers with insufficient sleep, vision problems and intoxication during nighttime leading to higher numbers of accidents at work zones; (4) adverse public reactions due to construction noise during nighttime; (5) difficulty in recruiting personnel in spite of the wage premiums that compensate for nighttime work; (6) difficulties in material delivery, utility services and urgent equipment repairs during nighttime hours (Shepard and

Cottrell, 1985); and (7) increase in cost for nighttime operations due to labor premiums and overtime, additional traffic control devices, additional artificial lighting arrangements, and higher engineering inspection costs (Hinze and Carlisle, 1990)

To overcome many of the above disadvantages, proper and adequate lighting arrangements need to be provided on nighttime construction sites. Lighting was reported to be one of the most important factors affecting quality, safety, cost and productivity of nighttime construction projects (Kumar, 1994). The design of lighting arrangements needs to be performed in a systematic and optimal way to achieve the best use of available lighting equipment. The dynamic nature of nighttime highway construction and maintenance projects in terms of variability of work zone locations and layouts within the same project requires an automated decision support system (DSS) that copes with these dynamic design aspects.

This paper presents an automated decision support system for the design of temporary lighting arrangements in nighttime highway construction operations. The system provides support for highway contractors and resident engineers in optimizing lighting design for nighttime construction. It is designed and developed to be (1) effective in providing near optimal solutions to the lighting arrangements; and (2) efficient in generating the required design in a reasonable time and effort due to the temporary nature of such lighting arrangements.

2. PROPOSED DECISION SUPPORT SYSTEM

The development of the system is attained through three main stages: 1) determining the design variables and constraints pertinent to the nighttime construction lighting arrangements; 2) identifying the objectives of the lighting design in highway nighttime construction and formulating them in a robust optimization model; and 3) implementing the optimization model using a multiobjective evolutionary algorithm.

2.1 Decision Variables

The following decision variables were identified for the lighting design in nighttime construction:

- Lighting equipment selection: The designer needs to choose from available alternatives of (a) ground mounted towers, (b) trailer mounted towers, and/or (c) equipment mounted luminaires.
- Types of luminaires: The type of lamp needs to be selected from available alternatives of: (a) metal halide lamps; (b) high pressure sodium vapor lamps; (c) halogen lamps and (d) low pressure sodium vapor lamps.
- Lamp lumen output: It represents the energy emitted from the lamp and influence visual comfort and illuminance (IESNA, 1998).
- Mounting height: It represents the vertical distance between the center of the luminaires and the pavement surface. Portable lighting towers are typically manufactured with adjustable mounting heights that can reach up to 25m.
- Lighting towers positioning: This variable represents the location of the lighting towers in the work zone. Lighting positioning affects the average illuminance and the uniformity of lighting in the work zone.
- Aiming angle of luminaires: It is the angle between the center of the luminaires beam spread and the nadir. This variable determines the directional distribution of lighting and affects the coverage area as well as the glare produced by the luminaires.
- Lighting tower rotations: This variable represents the rotation of the lighting tower luminaires around a vertical axis, which is needed, as a decision variable when the luminaires light distribution is not symmetrical. A proper rotation angle enables the designer to direct the lighting intensity towards the intended area and to minimize the lighting spillage to unnecessary directions, reducing light trespass that is a common source of complaints in nighttime construction in urban areas.

2.2 Lighting Design Objectives

The design of lighting arrangement in nighttime construction operations should satisfy the following objectives:

- Illuminance: The lighting system needs to maximize the average illuminance level in the construction work zone. An objective function was formulated using the point-by-point method to calculate the average horizontal illuminance in a grid of uniformly distributed points covering the work zone area. The horizontal illuminance at each point of the grid is calculated using the inverse square law (Pritchard, 1995) considering all light sources in the work zone.
- Uniformity ratio: The uniformity ratio needs to be minimized in order to ensure that light evenly reaches all areas in the work zone. This value is computed by dividing the average horizontal illuminance value over the minimum illuminance computed at any grid point in the work zone.
- Glare: Glare needs to be minimized in order to limit the visual impairments and/or discomfort experienced by the traveling public and workers. The veiling luminance ratio is used in road lighting as a control measure of glare (IESNA, 2000). The veiling luminance calculations in this model are formulated by adopting the same standard conditions for observer's sight direction and angles that are used in roadway lighting design due to the similarity of both cases. The observer's eye height is taken to be 1.45m, and the observation direction is the drivers' sight direction, which is parallel to the centerline of the roadway (IESNA, 2000).
- Lighting cost: The cost of a lighting system can be reduced by minimizing two major cost items: (a) ownership cost of the lighting equipment, which is either the cost to buy, rent, or lease the lighting equipment, and (b) operational cost of the lighting equipment, which is a function of the energy consumption of the lighting equipment.

2.3 System Implementation

The proposed optimization model was implemented using NSGA-II, which is an improved version of NSGA, referred to as the fast elitist nondominated sorted genetic algorithm (Deb *et. al.*, 2000). NSGA-II is a pareto-based approach that handles multiobjective optimization problems through the nondomination concept.

Many engineering design problems involve multiple objectives. Typically, engineering design problems have at least two objectives: the cost of the designed system that should be minimized, and a certain quality characteristics or utility from the system which should be maximized. While the single objective design formulation finds the best possible design solution that corresponds to the minimum or maximum value of the objective function, there is no single best solution in the multiobjective design. Rather we have a set of trade-off solutions generally known as the pareto optimal solutions or nondominated solutions (Deb *et. al.*, 2000).

Genetic algorithms (GAs) are search and optimization tool inspired by the mechanics of natural selection and genetics. Those algorithms adopt the survival of the fittest, and the structured exchange of genetic materials among population members over successive generations as a basic mechanism for the search process (Goldberg, 1989). Since its development, GAs have been widely used in various disciplines including engineering, science, business, and medicine (Chambers, 2001). The success of GAs in these fields can be attributed to their broad applicability in terms of their ability to handle various types of functions and to find global near optimal solutions in a multimodal search space (Deb, 1999).

Several approaches can be used to handle multiobjective optimization problems including: (1) weighted sum method; (2) goal programming approach; (3) constraint method; (4) lexicographic ordering method; (5) game theory approach; (6) gender-based GA; (7) multiple objective GA; and (8) non-dominated sorted GA

(NSGA, NSGA-II) (Coello, 1999; Deb *et al.*, 2000).

The proposed optimization model was implemented using NSGA-II due to its superiority over other multiobjective optimization tools in (1) providing the entire pareto optimal front of nondominated solutions in a single run; and (2) handling any number of objectives (Deb *et al.*, 2000). NSGA-II works by sorting the population members at each generation into sets of nondominated pareto optimal fronts. All points in each of the pareto optimal fronts are given the same fitness value according to their rank and solutions with higher rank have higher probability of being selected for reproduction (Deb *et al.*, 2000).

The optimization model was coded using C++ language to enable the evaluation of the formulated objective functions for a given set of decision variables. As shown in Figure 1, four functions were developed to calculate illuminance, uniformity ratio, glare, and cost to enable their evaluation in the optimization model. The photometric characteristics of the lighting tower luminaires (light distribution and lamp lumen output, and the reflectance characteristics of the pavement surface are entered as data files that can be accessed by these functions. Other decision variables such as number and position of lighting equipment, luminaires aiming angle, rotation, and mounting height are randomly generated for the initial population.

This initial population evolves over a number of specified generations in order to obtain a number of feasible solutions that are considered to be nondominated. The decision maker can select one solution for implementation from these sets to satisfy the particular design problem at hand.

3. APPLICATION EXAMPLE

An application example of lighting design for a work zone with a length of 27 meters, and a width of 10 meters is analyzed to illustrate the use of the proposed system. The design criteria in this example are specified to be: (1) a minimum average illuminance level of 100 lux to provide acceptable visibility for the construction activities in the work zone; (2) a maximum average illuminance of 200 lux to avoid light trespass to adjacent properties; (3) a maximum allowed uniformity ratio of 6; and (4) a maximum allowed glare (veiling luminance ratio) of 0.4.

Several runs were performed with different genetic algorithm parameters to study their effect on the convergence characteristics of the model. The output of the model in each run was

a set of nondominated solutions that satisfy the earlier described four design objectives. The solution space is a four-dimensional one, which makes the pareto optimal front for all functions impossible to view simultaneously. Two-dimensional slices from the pareto optimal front are therefore obtained to visualize the trade-offs between the different objectives. These slices present subsets of the nondominated solutions that represent the pareto optimal fronts considering two objective functions at a time. Six fronts are obtained for each run. For example, Figure 2 shows the trade-off between illuminance and glare/veiling luminance ratio objectives.

It can be recognized from the nondominated fronts that both objectives are conflicting, and an increase in the fitness of one function will lead to a reduction in the fitness values of the other. The shape of the nondominated fronts in terms of slopes provides insight on the nature of the trade-offs between different objectives. The designer can select a solution that produces low glare while making only a small compromise in illuminance.

Almost all fronts obtained from different runs have roughly the same shape, as shown in Figure 3. These fronts however have different properties in terms of their spread, and the values of the objective functions. It was also observed that the population size has a more important role in determining the quality of the obtained front than the number of generations as shown in Figure 3. The results obtained from the run that has a population size of 250 and 50 generations are better than the results obtained from the run that has a population size of 50 and 250 generations because the nondominated fronts in the former case have a better spread over the entire front, and also have better fitness values.

From the above analysis, one can see that it is possible to reduce the computational requirements of the lighting model, if the accompanying reduction in the fitness values of the objective functions is acceptable. It seems that the run with population size of 250 and 50 generations gave acceptable results with respect to the best run with a population size of 800 and 234 generations. The number of function evaluations in the former case with a reduced population size and reduced number of generations is less than 7% of the number of function evaluations in the run with a population size of 800 and 234 generations.

4. CONCLUSIONS

The results presented in this paper illustrates the capability of the proposed decision support system in: (1) handling multiobjectives in lighting design process simultaneously; (2) providing feasible solutions by satisfying the design criteria; (3) achieving good quality in lighting design rather than accepting the minimum requirements; (4) quantifying glare which is a major source of complaints in highway nighttime construction; (5) incorporating cost as an important objective in the optimization of the lighting design process; and (6) providing an automated and practical tool for the nighttime construction operations personnel to deal with the dynamic lighting design process.

5. REFERENCES

1. Chambers, L. (Editor) (2001) "The Practical Handbook of Genetic Algorithms Applications". 2001. Chapman and Hall/CRD. Florida.
2. Coello, C. (1999) "A Comprehensive Survey of Evolutionary-Based Multi objective Optimization Techniques". Laboratorio Nacional de Informatica Avanzada. Veracruz, Mexico. http://www.lania.mx/~ccoello
3. Deb, Kalyanmoy *et. al.* (2000) "A Fast Elitist Non- dominated sorting genetic algorithm for multi-objective optimization: NSGA-II". Proceedings of the Parallel Problem Solving from Nature VI Conference, 16-20 September, Paris.
4. Deb, Kalyanmoy. (1999) "Evolutionary Algorithms for Multi-Criterion Optimization in Engineering Design". Proceedings of Evolutionary Algorithms in Engineering and Computer Science (EUROGEN-99).
5. Goldberg, D. (1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley Longman, Inc.
6. Hancher, D., and Taylor, T. (2001) "Night-Time Construction Issues" Paper presented at the Transportation Research Board 80th Annual Meeting, January 7-11.
7. Hinze, J., and Carlisle, D. (1990) "Variables Impacted by Nighttime Construction Projects", *Final Report TNW 90-07*, Transportation Northwest, University of Washington, Seattle.
8. IESNA (2000) "IESNA Practice for Roadway Lighting". American National Standard / Illuminating Engineering Society of North America ANSI/IESNA RP-8-00. 2000 USA.
9. IESNA (1998)."IESNA Guide for Choosing Light Sources for General Lighting". Report DG-10-98. Illumination Engineering Society of North America 1998. New York.
10. Kumar, A. (1994) "Development of a Model for Determining Work Zone Illumination Requirements During Nighttime Highway Construction " PH.D Dissertation University of Florida.
11. McCall, H. C. (1999). "Report on the Department of Transportation's Administration of the Nighttime Construction Program", Report 98-S-50, State of New York Office of the State Comptroller.
12. Pritchard, D.C (1995) Lighting, 5th Ed., Longman Scientific and Technical, Burnt Mill, Harlow, Essex.

13. Price, D. A. (1986) " Nighttime Paving", Implementation Report CDOH/DTP/R-86/6, Colorado Dept. of Highways, Denver, Federal Highway Administration, April. 1986.

14. Shepard, F.D. and Cottrel, B (1985) "Benefits And Safety Impact of Night Work-Zone Activities", Report No. FHWA/RD-85/067, U.S. Department of Transportation, Federal Highway Administration, June 1985.

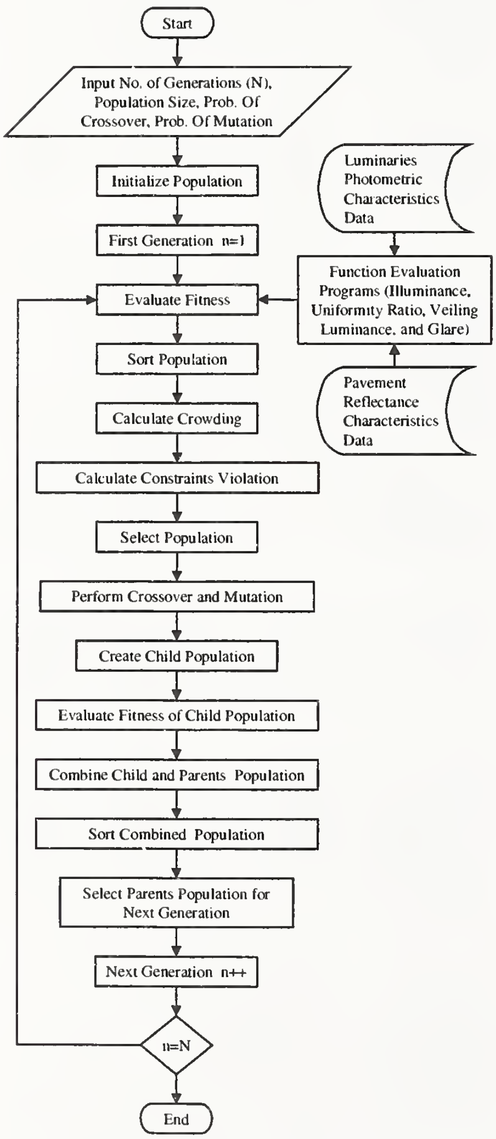


Figure 1. Lighting Design and Optimization Model

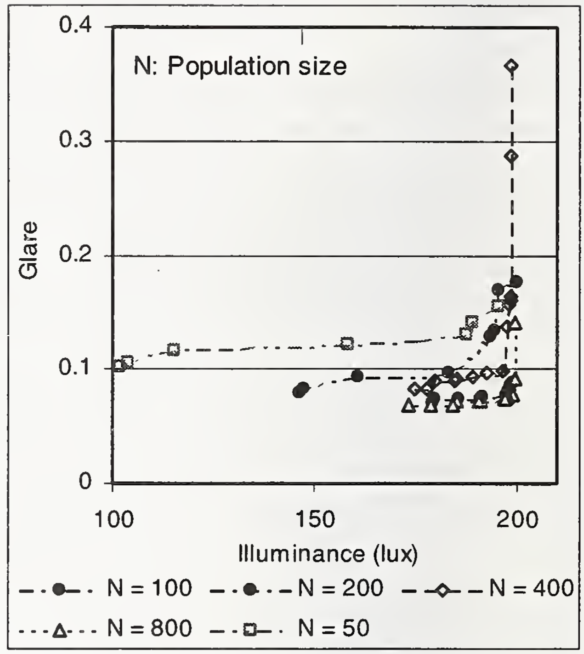


Figure 2. Illuminance / glare trade-off

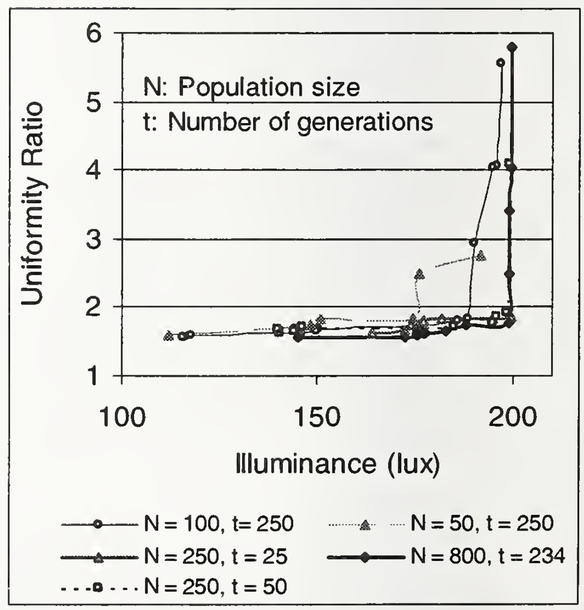


Figure 3. Illuminance / uniformity ratio trade-off

SESSION 3

DESIGN PRACTICES TO FACILITATE CONSTRUCTION AUTOMATION

De Bolder, an unusual example of off-site construction

by

Ger Maas, Bert van Eekelen[1]

ABSTRACT:

De Bolder ("The Bollard") is a 42.5 metre-high building with a circular cross-section of 30 metres and a weight of 25,000 KN. It was built in an industrial plant, transported a considerable distance across water, subsequently put ashore and placed on a foundation.

This study focuses on the differences in construction methods and the consequences of these differences.

Research aim: How can the De Bolder study results be used to improve traditional construction techniques?

Research questions: What are the differences between construction of De Bolder and construction on a traditional site? Where do these differences stem from? To what extent do these differences impact the actual construction period?

Results: use of this method depends on the characteristics of the object (volume, design and weight) and circumstantial attributes (transportability). The dynamics of transport determine the building's design; planning of an off-site construction process entails other dependencies, such as blurring of the distinction between structural works and finishing works.

KEY WORDS:

Building systems, industrialised construction, integrated product and process design, transport of building.

1. INTRODUCTION:

1.1 The initiators

This is a project involving the construction of a new Headquarters for the Mammoet Van Seumerengroep in Schiedam, the Netherlands, an organisation known throughout the world for its unique achievements in the offshore industry and the associated heavy transport operations. In 2001, the company successfully raised the Russian submarine Kursk, an operation considered impossible by many.

1.2 The design

The building was to be designed in the form of a bollard. It was to be a compact, tenstorey construction with a cylinder-shaped building volume, which, by reason of its constricted shape, resembled a bollard. The building is situated on the axis of the Benelux tunnel, the route of the A4 motorway under the river Meuse (the Rotterdam harbour), which makes it visible from a long way off by the 150,000 motorists who travel along this motorway each day.

[1] Eindhoven University of Technology, Faculty of Building, Architecture and Planning, Postbus 513, NL 5600 MB Eindhoven, The Netherlands
g.j.maas@bwk.tue.nl

The building's compact bollard shape made it ideal for transporting. Because of the production capacity at Grootint Zwijndrecht – one of the subsidiaries of the Seumerengroep where oil platforms are made – the idea was conceived of constructing this building in the industrial plant in Zwijndrecht and transporting it by water to its ultimate location in Schiedam.

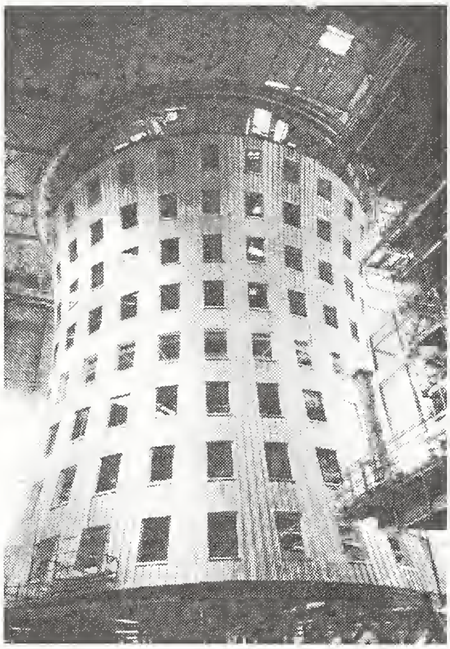


Figure 1 De Bolder under construction in the industrial plant

The project has the following characteristics: diameter: approx. 30 metres; height: 42.5 metres; mass: 25,000 KN; transport distance by water: 25 km; over land: 300 metres; method of transport: self-driven platform trailers and seagoing pontoons; largest obstacle: the Botlekbrug bridge with a maximum headroom of 45.7 metres at lowest possible water level; construction period (superstructure): mid-July 2001 to 29 January 2002.

As there was no experience with the prefabrication of this type of office building, the decision was taken to develop this innovative industrialised construction method parallel to a building to be constructed using traditional techniques.

This meant that it had to be possible to build the structure along traditional lines and also in such a way that it could be transported. This applied to all aspects of the construction, i.e. functional and spatial development, intended constructional and installation technologies, as well as other specific structural details and dimensioning.

As a result, the building differs only slightly from traditional structures, with the exception of a few components. It may be assumed that it is possible to optimise the construction yet further if the transportable variant is opted for at an earlier stage. On the other hand, limiting the project by allowing for both options did result in an explicit focus on specific aspects such as planning effects and specific circumstances on the construction site, without these aspects being related to a very different building concept. To all intents and purposes, De Bolder is a traditional structure, but one built under industrialised circumstances in a factory setting, and one that could be transported.

2. RESEARCH QUESTIONS

- What are the differences between construction of De Bolder and construction on a traditional site?
- Where do these differences stem from?
- To what extent do these differences impact construction time?
- How do the separate building activities relate to one another?
- How can the results of the research into this type of conditioned construction work be used to improve current construction processes?

3. RESEARCH METHOD

3.1 Research limitations

Research work was carried out between November 2001 and June 2002.

When the study started, the project was already at its final completion stage. Apart from the fact that the outside walls had not yet been sealed off, it was not immediately possible to see that this building had been factory-built. Nor was it possible during construction to identify and measure productivity or failure costs and to compare these with standard times or costs of traditional structures cited in the relevant literature.

3.2 Research approach

This study was conducted by reconstructing the preparation and construction processes as faithfully as possible. By collecting data and conducting interviews with people directly involved in the construction work, it was possible to paint a picture of the way in which preparations and actual construction work had been carried out. Of particular interest were the decisions taken and the criteria and considerations leading up to them. This reconstruction was compared to the common Dutch construction methods. Traditionally prefabrication is limited to certain sections of a building.

4. THE RESULTS

4.1 Planning

4.1.1 Introduction

An important difference between traditional techniques and those used in the construction of De Bolder is the drastic reduction in building time. Despite the time required for transport, a net time gain of 22 weeks was made. The difference in construction time is made up as follows:

1. Work under industrial circumstances (- 6 weeks)
2. Changes in planning (order and sequence of building activities) (-20 weeks)
3. Preparations for transport, transporting and unloading (+ 4 weeks)

4.1.2 Work under industrial circumstances

The fact that works could continue irrespective of the weather resulted in a direct time gain. A standard of 29 days on which work is held up on account of weather conditions applies to this region of the Netherlands for the duration of any construction work. This downtime includes seven working days due to frost, nine due to rain and eleven due to wind.

Although deviations will, in practice, occur as regards actual weather conditions and there will be workable days on which weather does hold up work but only to a certain extent, these 29 days are a generally observed rule. Eliminating downtime not only shortens the construction period but also does away with the unpredictability of work during the winter. Work can be carried out under all circumstances, irrespective of the weather, so that the work itself and the link between the project as a whole and the various individual activities involved (sequence and drying times, for instance) can be far more strictly planned.

4.1.3 Changes in planning

If we compare the planning work of this construction project with those of traditional projects of the same type and scope, three aspects among the various activities stand out most.

Firstly, the erection of the supporting structure can start without prior work on the foundations. The floors in industrial plants where oil platforms and offshore installations are built have sufficient load-bearing capacity to support a building of this size. The building of the eventual foundation for the building can run parallel to the construction of the upper structure at the final location, generating a time gain of eight weeks in this case.

A second difference is the direct succession of structural work and finishing work. At the completion stage, materials are used that are susceptible to weathering. The building has to be fully glazed before most of the finishing work can be started, which is why it is usually not worthwhile starting finishing work immediately after the structural work. In the case of De Bolder, the entire project was both water and wind ‘resistant’ from the onset and conditioned construction was, therefore, possible. This allows for a direct start of the finishing work immediately following the structural work, storey by storey (by tradition, the building was built up from the bottom). Although, as far as safety was concerned, various incompatible activities were being conducted at the same time, conditioned circumstances on the construction site made it possible to keep this under control by working with two separate construction teams in different shifts throughout the critical period.

The third difference lies in the timely installation of costly equipment. As a rule, the building should be appropriately sealed off before costly equipment such as air conditioning and telecommunication systems and ICT networks are installed. Especially users equipment is often put off until after final acceptance. The result is that shafts and ceilings have to be reopened. In the case of De Bolder, facilities such as those mentioned above were assembled and installed in good time, seeing as the risk of theft or vandalism was virtually eliminated. The above two differences resulted in a time gain of twelve weeks; this does not include the time gained after completion once the end users started moving in. The total time gain of the three mentioned differences is 8 + 12 = 20 months.

4.1.4 Productivity

Finally, we must mention a further key factor, which we have not been able to explore in greater detail: actual productivity during construction work. Various factors play a part, for instance the extent to which activities can be planned, whether they can be physically prepared, the logistic advantages of an industrial setting (both material advantages and advantages as regards the information flow), and matters relating to motivation of the employees on the job.

Previous analyses have shown that improvements to circumstances on the construction site can lead to a dramatic increase in the productivity of construction workers, from 50% to 75% of the available working hours. This may be even more in an industrial production setting such as the one used for De Bolder. On the other hand, the consequence of adapting this kind of building structure to an industrial setting in which high safety requirements and stricter procedures apply for working on offshore installations may also have a decelerating effect.

We recommend that further case studies investigate productivity more thoroughly, as this can lead to new findings of practical use to more traditional building methods.

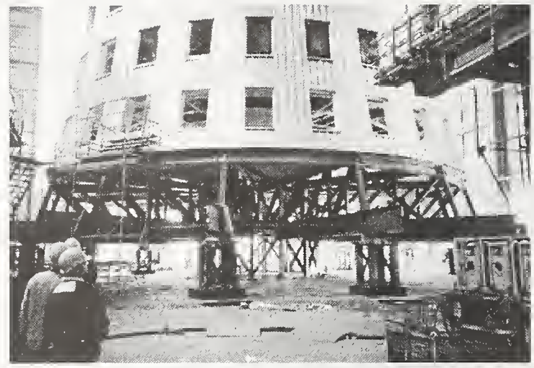


Figure 2 De Bolder on the transition structure

4.2 Design and choice of materials

Design requirements, weight and deformations have to be taken into account to make the transport of a structure possible, in fact the specific transport modalities have to be taken into account as early as the initial design stage. Vehicles bear the weight of the structure during transport; in the case of De Bolder these were what are known as self-steering platform trailers commonly used in the offshore industry.

Complete oil platforms and complex installations are built in factory sheds and transported to the wharf on such vehicles. This was also done with De Bolder. The number of vehicles is, in theory, infinite – the weight to be carried is limited only by the axle load-bearing capacity.

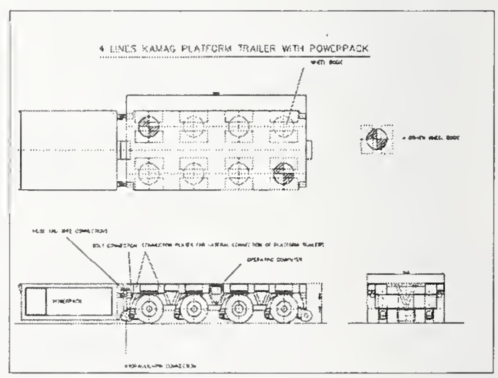


Figure 3 The platform trailers

The ground plan of De Bolder specified a maximum weight of 2,5000 KN. This meant that, compared to the usual construction method, the weight of this building with its nine office levels had to be reduced. Monolithic inner leaves and floors are generally used in buildings such as these, the decisive factors being ease of realisation as well as cost reduction. By opting for more expensive composite floor slabs and inner leaves with a wood frame construction, it was possible to reduce the weight of the building by 20%.

Design factors are likewise important, which is why the office-dividing partitions were not erected until after transport. Although weight can be distributed across several platform trailers, the structure should be such that it is possible to transport it without running the risk of major deformations. For the offshore industry this means that the load is always grouped in such a way that three ultimate bearings are created, so that deformations and leaning during transport can be checked and managed.

Traditional structures always transfer their own weight and utilisation load to the foundation construction, following a logical grid. In the case of a building that is to be transported, that load has first to be redirected to three bearings and then redistributed over the several axles.

In practice, therefore, this means that a transition structure is required for the substructure of the building, something not usually found in traditional structures.The transition construction can be part of the building's own structural framework or made as a temporary structure; the latter was opted for in the case of De Bolder.

It is, in any case, evident that both the transportable designs of the structure and the structural features this requires have to be taken into account as early as the initial design stage.

4.3 Costs

De Bolder has been built at a cost akin to that of a traditional building. This does require some comment, however. There was access to an industrial plant of sufficient size and load-bearing capacity and with all the necessary amenities. Transport facilities were also available and hardly any equipment needed to be bought specially for this project.

Given the overall cost of the project, the transport costs were by no means the heaviest. These amounted to EUR 113,250 i.e. the sum of the costs for the platform trailers plus the seagoing pontoon.

The costs for the temporary load-bearing construction, i.e. the transition construction that made it possible to carry and transport the building on the platform trailers, were significantly higher, viz. EUR 1,250,000. In addition, the wharf at the building's final destination had to be reinforced to receive the platform trailers, the costs for which came to EUR 50,700.

These excess costs were eventually compensated by the savings made on construction time, the reduced risk of theft and wilful destruction, as well as other economies.

As both the industrial plant and the final location are by the waterside, making transport over land difficult, we investigated to what extent circumstantial factors have an impact on the transportability of such buildings. Using an abridged sensitivity analysis, we found that transport distances, be they by water or over land, had no profound effect, provided no cost-raising obstacles had to be overcome. Given the flatness of the land and the abundance of waterways in the Netherlands, this concept opens up possibilities for future use.

4.4 Social effects

Cutting back construction times has been an issue in the Netherlands for some time now. This can be achieved in a number of ways. Optimising conditions on construction sites is in itself one way of improving the predictability and planability of construction work, thus also avoiding interruptions, loss and downtime. In that sense, much can learnt from the factory-based approach used in the construction of De Bolder. However, there is no need to opt for total prefabrication or transport of an entire structure.

Yet it would seem that in the Netherlands there is an increasing shift in the characteristics of the surroundings in which construction work is carried out, from erecting new buildings in 'green field' sites, i.e. within perimeters, to embedding construction works in existing urban settings, adapted to existing daily operational activities. Such construction projects are found in the ever more densely built-up inner cities, near airports, railway stations and large complexes such as exhibition sites, educational institutes and health care institutes. Studies of the effect that the surroundings have on building work in 'revitalisation' areas have shown that disruptions should be kept to a minimum. Not only that, the nature and duration of such inconveniences, as well as their predictability are also crucial. The traditional construction process 'within the perimeters' is not geared to that and numerous practices and sector-specific improvements in traditional building projects are inconsistent with environmental requirements.

It is for that reason in particular that the De Bolder project has attracted so much attention from the construction industry as a whole.

5. CONCLUSIONS/LESSONS:

- If a design is suited for off-site production and transport, there are many advantages to be gained from alternative construction process dependencies resulting in disruption-free building with shorter interfaces and a blurring of the distinction between structural work and finishing work.
- The success of De Bolder has established the efficacy of this technology; its application depends on the volume, design and weight (object attributes) of the structure and its transportation capacity (circumstantial attributes).
- The design of such structures depends largely on dynamic aspects (unusual for buildings), which implies that there are other construction problems that need to be addressed at an earlier stage in the development process.

6. REFERENCES

1. M.R.Tuinenburg (484176) TU Eindhoven, Techniek en Maatschappij, TIB AE, faculteit Technologie Management: Transportsystemen voor grote bouwwerken, een casestudy naar alternatieve weg- en watertransportsystemen voor "De Bolder" juni 2002

2. http://www.staalbouw.nl/invulpaginas/dnst_proj_bolder.html

3. http://www.mammoet.com/projects/index.htm

4. http://www.heerema.com/content/hfg/yards/grootint/projects/bolder.htm

AUTOMATED DATA ACQUISITION AND PLANNING OF HIGHWAY CONSTRUCTION

by

Ahmed Hassanein[1] and Osama Moselhi[2]

[1] *PhD Candidate, Dept. of Building, Civil and Environmental Engineering*
Concordia University, Montréal, Québec
hass@alcor.concordia.ca

[2] *Prof. and Chair, Dept. of Building, Civil and Environmental Engineering*
Concordia University, Montréal, Québec
moselhi@cbs-engr.concordia.ca

ABSTRACT: This paper briefly describes a model, designed to automate data acquisition and analysis for planning and scheduling highway construction projects. The paper briefly describes the proposed model, and focuses primarily on its automation aspects. Geographic information systems is employed to analyze spatial data and estimate cut and fill quantities. The model also stores activities typically involved in highway construction and automatically generates the precedence network respecting job logic. The model generates digital terrain models to represent the ground topography and underlying soil strata. Transition of soil strata between borehole locations can either be: 1) automatically generated by the model; or 2) defined by the user. The model employs the mass haul diagram to develop the optimum earthmoving plan. It accounts for the presence of transverse obstructions, such as rivers and existing highways, when developing the earthmoving plan and defining project activities. The model is implemented in a prototype software that operates in Microsoft Windows environment. It provides a user-friendly interface, including menus, dialog windows, and graphical capabilities to expedite data input and retrieval. Several input and output formats are accepted to facilitate data sharing with commercially-available software packages.

Keywords: automation; earthmoving; geographic information systems; highway planning

1. INTRODUCTION

Although highway construction projects are typically large in terms of capital requirements (Adeli and Karim 1997), little effort was made to automate the planning stage of this class of projects. The effort required to develop a competent plan is perceived as a major obstacle to developing high quality schedules (Chevallier and Russell 2001). Earlier studies (eg. Herbsman 1987; NCHRP 2000) revealed the dissatisfaction of parties involved in planning and scheduling highway operations with available planning and scheduling tools. Of these operations, earthmoving operations generally represent a major bid item, making their optimization of paramount importance to developing successful bids (Jayawardane and Harris 1990). Acquiring and analysing spatial data has been perceived as a time-consuming process, which is susceptible to certain biases (Fan *et al* 2001). Several methods have been proposed for accurate estimation of cut quantities (e.g. Siyam 1987; Epps and Corey 1990; Easa 1992). On the other hand, considerable work has been done to optimize earthmoving operations (e.g. Stark and Mayer 1983). However, none of these models: 1) provide for automated data entry; 2) account for the impact of transverse natural and man-made obstructions on the earthmoving plan; 3) support three-dimensional visualisation; or 4) provide scheduling capabilities.

2. PROPOSED MODEL

This paper briefly describes a model designed to automate the above-mentioned functions, and focuses primarily on the automation aspects. Transverse obstructions, such as rivers are considered when developing the work-breakdown structure (WBS). The model stores a list of activities typically encountered in highway construction, and generates the precedence network respecting the job logic. Geographic information systems (GIS) is employed to acquire and analyze spatial data and generate the optimum earthmoving plan using the mass haul diagram. The model automatically generates digital terrain models (DTMs) to represent the original ground topography and underlying soil strata. When estimating quantities of cut and fill, the model accounts for variations in swell and shrinkage factors of soils of these strata. The model is implemented in a prototype software that operates in Microsoft Windows environment. It accepts several input formats and provides graphical and tabular reports.

The model is schematically illustrated in Figure 1. It utilises GIS to acquire and analyze spatial data. A knowledge base is employed to define the WBS and generate the precedence network respecting the job logic. Relational database models store: 1) swell and shrinkage factors of common soil types; and 2) available resources. Once the WBS and precedence network have been defined and crews have been assigned to their respective activities, the project can be scheduled. In view of the highly repetitive nature of highway construction, the scheduling engine employs resource-driven scheduling. As the figure shows, the model generates tabular and graphical reports, and generates the optimum earthmoving plan using the mass haul diagram. The main functions of the model are discussed below.

2.1. Data Acquisition and Analysis

Figure 2 illustrates the sequence of operations to automate data acquisition and analysis. The model extends the capabilities of the traditional mass haul diagram (Stark and Mayer 1983) to account for: 1) the presence of different soil strata in cut sections; 2) the presence of transverse natural and/or man-made obstructions; and 3) varying degrees of soil compaction. The model generates three-dimensional DTMs to represent the original ground topography and underlying soil strata. These DTMs are stored as triangulated irregular networks (TINs) so as to reduce demand on the system's memory without sacrificing accuracy (Oloufa 1991). An iterative procedure is proposed to generate TIN nodes along contour lines (see Figure 3). Initially, the iteration number (R) and contour line identifier (L) are initialized to unity. Next, the node number (i) and distance from start of contour line (d_i) are initialised to unity and zero, respectively. If (R=1), the distance along the contour line to the next node, "D_{iRN}", is set equal to the horizontal component of the perpendicular distance to the nearest contour line. Otherwise (i.e. R>1), D_{iRN} is reduced to half its value in the previous iteration.

This process is repeated until nodes are generated along the entire length of the contour line (nested loop 1). The same procedure is applied to generate TIN nodes along remaining contour lines (nested loop 2) enabling the generation of a DTM. Next, the volume of soil contained under that DTM, "Vol_R", is computed. If (R=1) then "R" is incremented and the whole process is repeated (loop 1). Otherwise, estimated volumes of soil from the last two iterations are compared. If the change in "Vol_R" is less than the specified tolerance (default=15%), then the process is complete. Otherwise "R" is incremented and the process is repeated until the required accuracy is attained.

Strata profiles can either be: 1) defined by the user, by specifying the strata connectivity between boreholes; or 2) automatically generated by the model, and revised as needed by the user. In the latter case, each soil layer in a borehole is assigned a number, "i", that accounts for its sequence from top to bottom, starting with i=1 at ground level. The model assumes that the assigned number for each layer can go up or down by one level between two successive boreholes. If a stratum is not found in neighbouring borehole test results, it is assumed that the stratum tapers linearly to reach a thickness of zero (0) at the location of that borehole.

The volume of soil contained between two successive sections in the highway embankment is estimated knowing the end areas and the distance between them. A modification of the average-end-area method proposed by Epps and Corey (1990) is employed. The methodology proposed by Easa (1992) is employed to estimate the volume of soil contained in curved portions of the highway embankment. This methodology requires that end section be simplified, and the procedure proposed by Easa (1989) is employed for that purpose. The model automatically determines which sets of equations to employ when estimating cut and fill volumes.

2.2. Work-Breakdown Structure

Transverse obstructions, such as rivers and creeks, play a major role in performing the WBS (Alkass and Harris 1991). There are two types of obstructions: 1) surmountable, where access is granted across the obstruction at an overhead (time and cost); and 2) insurmountable, where no access is granted. Work zones are normally defined based on the locations of insurmountable obstructions, while segments are defined based on the locations of surmountable ones. The model stores a list of activities typically encountered in highway construction and divides activities into work packages based on the locations of transverse obstructions. The resulting WBS is shown in Figure 4.

2.3. Precedence Network

The model automatically generates precedence networks defining the construction of new highways, as well as the rehabilitation of existing ones. Generated precedence networks are generated detailing operations along the main project route as well as the construction of: 1) overpasses; 2) transverse overpasses; and 3) interchanges. The model enables the selection of any or all of the stored activities, and the precedence network is generated accordingly. The model also enables the definition of additional activities, along with their precedence relations.

3. COMPUTER IMPLEMENTATION

The proposed model is implemented in a prototype software that operates in Microsoft Windows. The software: 1) is user-friendly; 2) generates graphical and tabular output; and 3) enables data exchange with commercially available software. It provides an interface with dialog windows, menus, toolbars and a status bar. In view of its widespread acceptance, ArcView GIS (version 3.1) is employed as the GIS engine, while three-dimensional analysis is carried out using ArcView 3D Analyzer. Microsoft Access is employed as the database management system (DBMS). The software accepts a wide variety of input formats, including AutoCAD files, digitized blue prints and GPS data points. It provides three-dimensional visualization of: 1) final highway embankment; 2) ground topography; and 3) underlying soil strata.

4. SUMMARY AND CONCLUDING REMARKS

This paper briefly describes a model, designed to automate: 1) spatial data acquisition; 2) the generation of the optimum earthmoving plan; 3) the development of the project's work-breakdown structure; and 4) the development of precedence network respecting planned job logic. The paper focuses primarily on the automation aspects of the developed model, which employs geographic information systems to compute cut and fill quantities and generate the optimum earthmoving plan. A methodology is proposed to generate digital terrain models representing the original ground terrain and estimate quantities of cut and fill. A relational database stores swell and shrinkage factors for common soil types, enabling accurate estimation of soil quantities. The model accounts for the presence of transverse obstructions on the work breakdown structure and consequently the proposed earthmoving plan.

The model is implemented in a prototype software that operates in Microsoft Windows environment. The model demonstrates the potential of GIS as an emerging tool in acquiring and analyzing spatial data. Using the proposed model can significantly reduce the time and cost required to plan highway construction, while enhancing the accuracy of the developed estimates.

5. REFERENCES

Adeli, H. and Karim, A., 1997, "Scheduling/Cost Optimization and Neural Dynamics Model for Construction", Journal of Construction Engineering and Management, ASCE, Vol. 123, No. 4, pp. 450-458.

Alkass, S. and Harris, F., 1991, "Development of an Integrated System for Planning Earthwork Operations in Road Construction", Construction Management and Economics, E & FN Spon, Vol. 9, pp. 263-289.

Chevallier, N. and Russell, A.D., 2001, "Developing a Draft Schedule Using Templates and Rules", Journal of Construction Engineering and Management, ASCE, Vol. 127, No. 5, pp. 391-398.

Easa, S.M., 1989, "Simplifying Roadway Cross Sections Without Reducing Volume Accuracy", Canadian Journal of Civil Engineering, CSCE, Vol. 16, pp. 483-488.

Easa, S.M., 1992, "Estimating Earthwork Volumes of Curved Roadways: Mathematical Model", Journal of Transportation Engineering, ASCE, Vol. 118, No. 6, pp. 834-849.

Epps, J. and Corey, M., 1990, "Cut and Fill Calculations by Modified Average-End-Area Method", Journal of Transportation Engineering, ASCE, Vol. 116, No. 5, pp. 683-689.

Fan, L., Ho., C. and NG, V., 2001, "A Study of Quantity Surveyors' Ethical Behaviour", Construction Management and Economics, E & FN Spon, Vol. 19, pp. 19-36.

Herbsman, Z.J., 1987, "Evaluation of Scheduling Techniques for Highway Construction Projects", Transportation Research Record 1126, National Research Council, Washington, D.C., pp. 110-120.

Jayawardane, A. and Harris, F., 1990, "Further Development of Integer Programming in Earthwork Optimization", Journal of Construction Engineering and Management, ASCE, Vol. 116, No. 1, pp. 18-34.

NCHRP (National Cooperative Highway Research Program), 2000, *Project Management Information Systems*, Synthesis of Highway Practice #282, Transportation Research Board, National Research Council, National Academy Press, Washington, D.C.

Oloufa, A.A. 1991. "Triangulation Applications in Volume Calculation", Journal of Computing in Civil Engineering, ASCE, Vol. 5, No. 1, pp. 103-121.

Siyam, Y.M., 1987, "Precision in Cross-Sectional Area Calculations on Earthwork Determination", Journal of Surveying Engineering, ASCE, Vol. 113, No. 3, pp. 139-151.

Stark, R.M. and Mayer, R.H., 1983, *Quantitative Construction Management: Uses of Linear Optimization*, John Wiley & Sons, Inc, New York.

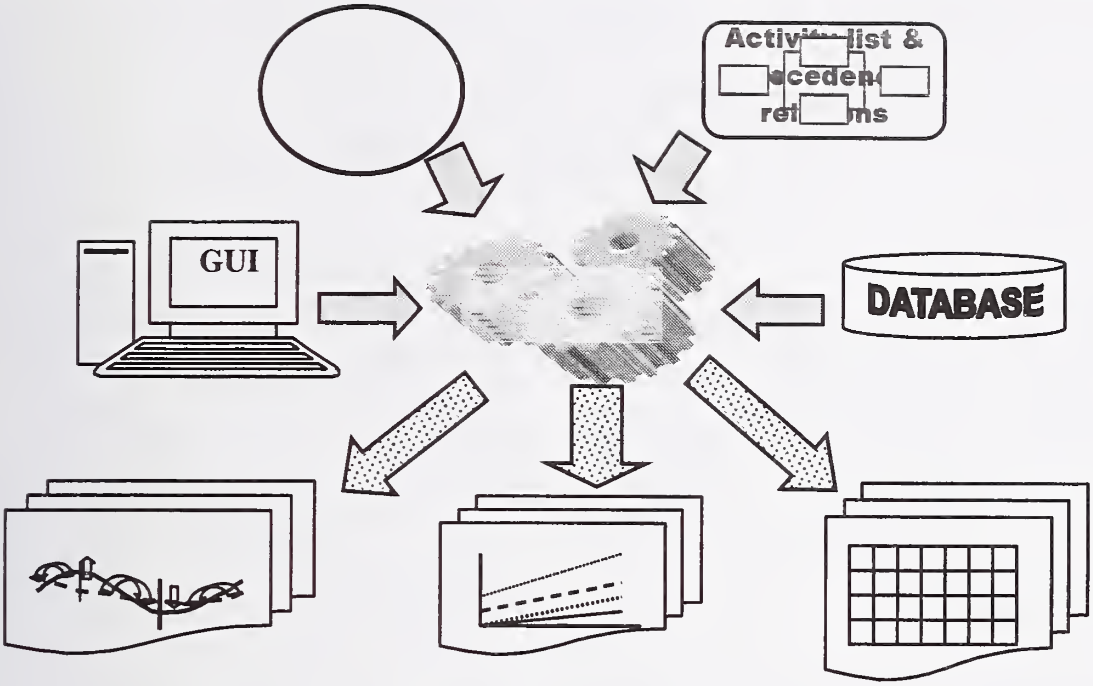


Figure 1: Proposed model

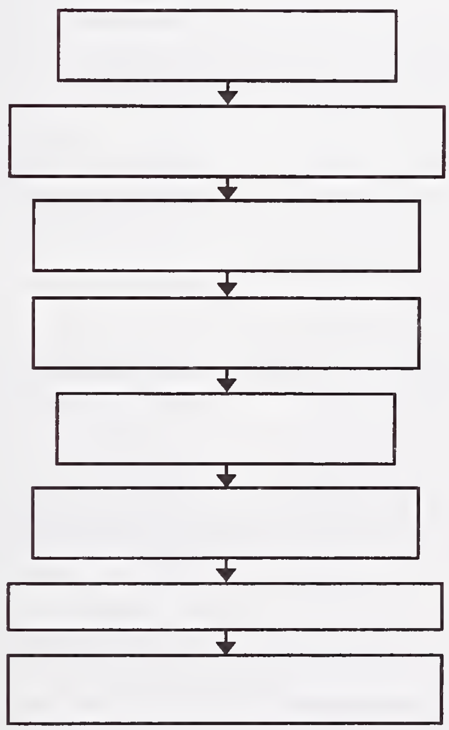


Figure 2: Procedure for analyzing spatial data

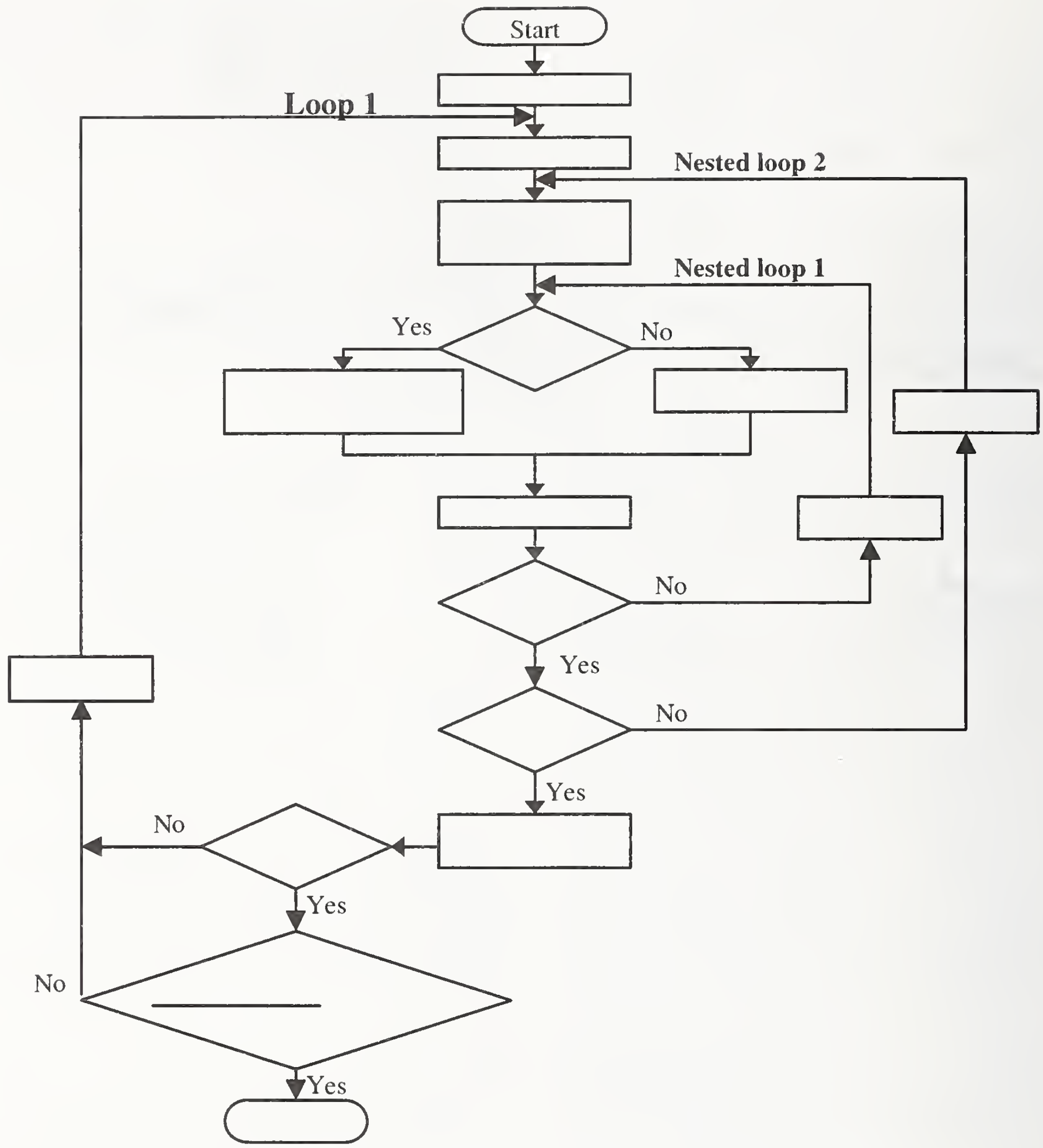


Figure 3: Developed algorithm to generate TIN nodes

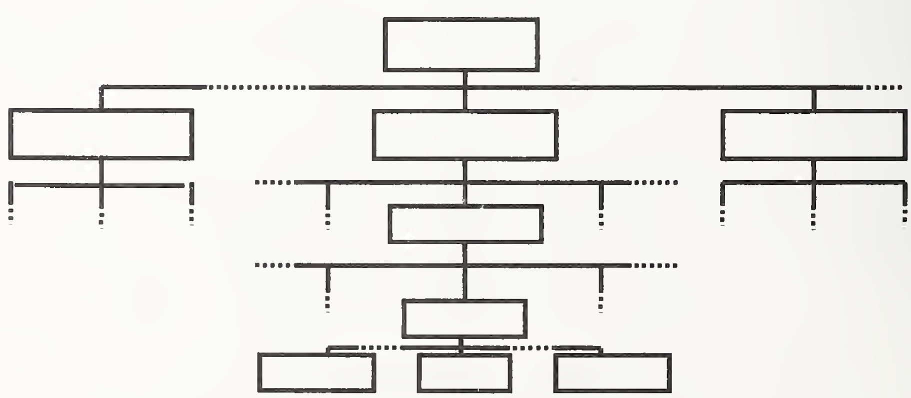


Figure 4: Proposed WBS

ASPHALT QUALITY PARAMETERS TRACABILITY USING ELECTRONIC TAGS AND GPS

François Peyret[1], Robert. Tasky[2]

ABSTRACT: This paper begins by introducing the developmental work carried out within the scope of the OSYRIS project, in the aim of proposing a new IT infrastructure for road-building sites and by drawing attention to the need for an electronic connection between paving operations and the asphalt mixing plant. The set of system specifications required to satisfy this connection is then given, along with a state-of-the-art survey. The central section of the paper presents the proposed solution, which is based on both electronic tagging of the trucks hauling asphalt and a GPS positioning of the asphalt pavers. The first experiment conducted in 2001 is also described and the paper closes with concluding observations and an outlook.

KEYWORDS: information and communication technologies, asphalt fabrication and laying, tracability, electronic tagging, GPS, quality control.

1 INTRODUCTION

Recent research initiatives, such as the European project OSYRIS, have been proposing new IT infrastructure in support of asphalt laying and compacting operations. OSYRIS focuses on the laying and compacting tasks but does not directly address the asphalt fabrication operation, which proves to be of greatest importance since this operation serves to set the quality of the asphalt material that will constitute the pavement.

Some quality assurance procedures have already been introduced to monitor and assess material quality at the plant, yet no direct connection between fabrication and paving operations has been established, hence no direct use can be made of quality assessment at the plant during paving and compaction processes. The aim of the present research is to create an electronic link in order to fill a major void in the sequence of electronic monitoring procedures for road asphalt sites.

2 THE OSYRIS PROJECT

2.1 Key Issues And Philosophy

OSYRIS ("Open System for Road Information Support") is a European Union-funded project devoted to developing information infrastructure for road construction and maintenance processes [1].

It is intended to enable both contractors and road owners to generate their own knowledge bases and quality assurance systems, which are now expected of them given that access to critical information in real time is one of the key challenges facing the new century.

The OSYRIS philosophy lies in building openness and modularity into a system through use of compliant components. OSYRIS data storage and management is based on a product model of the road specially designed to be: compatible with the latest road management databases, object-oriented, and located within a 3D geographical reference.

---

[1]*Site Robotics and Positioning Team, LCPC (Laboratoire Central des Ponts et Chaussées), Bouguenais, France, francois.peyret@lcpc.fr*

[2]*SEMR - CETE Normandie-Centre, Blois, France, robert.tasky@equipement.gouv.fr*

The OSYRIS project was launched in February 2000 and will last until the beginning of 2003. The Consortium partners consist of: the University of Karlsruhe (IMB) - Germany; LCPC - France; Moba - Germany; Tekla Corp. - Finland; and Skanska - Sweden. The project is funded under the European Union Program "Growth" (Contract No.: G1RD-CT1999-00080).

2.2 Expected outputs

Four following commercial OSYRIS products have been foreseen:

- *OSYRIS Standards*: a set of several pre-standards describing the format and contents of information exchanges taking place between OSYRIS components,
- *OSYRIS Machine*: a set of on-machine components developed to fulfil a contractor's basic needs in supporting paving and asphalt compaction work, for a typical asphalt-laying set-up of 1 paver and 2 rollers,
- *OSYRIS Machine Extensions*: a set of options to extend the basic *OSYRIS Machine* configuration,
- *OSYRIS Design and Documentation*: consists of a line of software for the planning, analysis, documentation and long-term storage of roadwork parameters. These products are intended primarily for consultants and road owners.

3 SPECIFICATIONS OF THE MATERIAL TRACABILITY SYSTEM

3.1 The Quality Control Chain

Generally speaking, material tracability is to be performed through eight main phases, as follows:

1. Supply of components (aggregates, binders, etc.) to the asphalt plant
2. Storage of components in piles or tanks
3. Transfer of components from storage devices to the mixing equipment
4. Fabrication, consisting of component grading and mixing, and constitution of material batches
5. Output of material batches from the plant
6. Reception of batches at the asphalt-laying site
7. Laying of the material
8. Compaction of the material.

This quality control chain contains a major natural discontinuity: the transport and delivery of materials from the plant to the construction site, between phases 5 and 6. This discontinuity often accounts for lost quantitative and qualitative information and, in some instances, delivery location errors occur. Consequently, it was decided that he system should be capable of:

- automatically collecting the most relevant fabrication-related information from the plant's existing monitoring and quality control systems for each batch of material;
- automatically downloading this information into a tag mounted on the truck headed for the paving site, along with the batch of material;
- automatically transferring this information once on site to the paver information system for use with respect to execution and documentation purposes;
- automatically locating and dating the information along the road at the particular place where the corresponding material has been laid.

4 DESCRIPTION OF THE PROTOTYPE SOLUTION

4.1 The Concept

The solution we have chosen is based on the two following technologies:

- "Radio Frequency Information Data" (RFID), for storing the data into electronic tags; and
- "Global Positioning System" (GPS), for positioning the material parameters with respect to the road-building project.

The data related to a batch of material are extracted and transferred to the tag from both the quality-monitoring computer (responsible for recording the quality parameters from the plant's main control computer) and the weighing table computer that measures the load of each truck as it leaves the plant. The transfer is performed through an interfacing box (called "i-box") connected to an antenna. Data are conveyed to the site by the tag; while the truck is loading asphalt into the paver hopper, the data are downloaded into the OSYRIS paver's computer by a similar transmission chain working in the opposite direction.

Figure 1 shows the physical structure and data flow of the system.

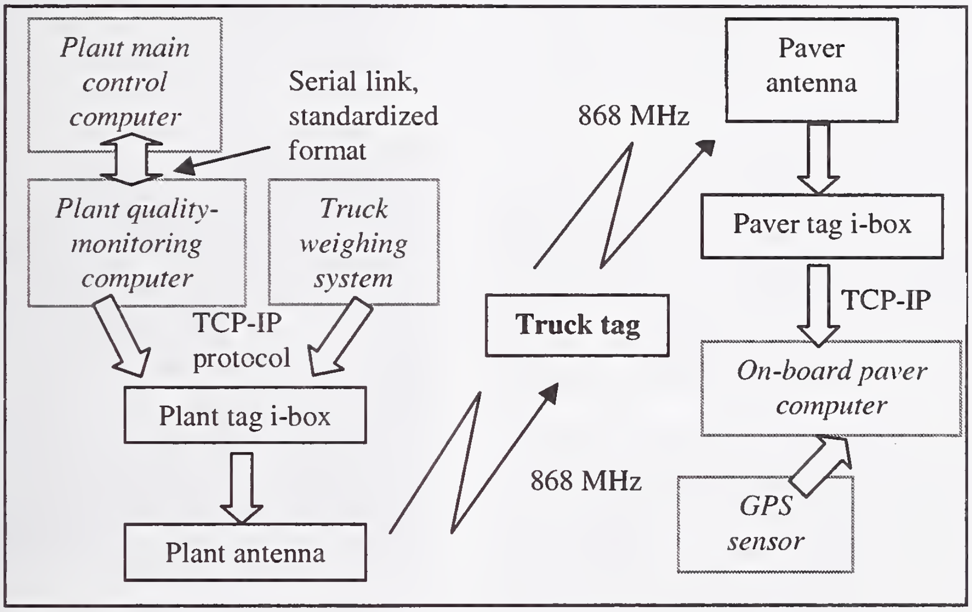


Figure 1: General structure and data flows of the system

For purposes of the feasibility experiment and due to the unavailability of the OSYRIS on-board computer and GPS receiver, SEMR and LCPC used a dedicated computer along with a GPS specific sensor. A customized software program has also been written for this demonstration.

4.2 Electronic Tagging System

The device chosen was "Flex ID", marketed by the German company DEISTER [2]. The main features of this system are as follows:

- transmission frequency: 868 MHz,
- tag dimensions: 15 cm x 3 cm x 2 cm,
- maximum number of antennae per i-box: 4,
- storage capacity of tags: 8 to 32 Kbytes,
- maximum read/write distance: nominal: 30 m; effective: 15 m,
- possible links between i-box and tag: TCP/IP, RS232 or RS485.

Figure 2 shows the tags and the antenna.

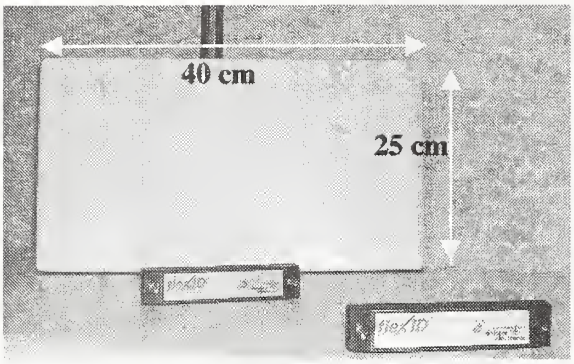


Figure 2: The Flex ID antenna and two of the tags

4.3 GPS Positioning

A stand-alone GPS is capable of providing 2D accuracy of approximately 3 to 5 meters since the Selective Availability (SA) feature was switched off by the USA in May 2000. This accuracy is not good enough for our needs, as we require being able to distinguish between the various lanes of a road.
For this reason, the GPS sensor we selected is an integrated DGPS sensor manufactured by the OMNISTAR Company; it is composed of an 8-channel Trimble receiver, a GPS antenna and a communications satellite antenna, to receive the differential GPS corrections, all embedded into the same housing.

4.4 On-board computer and software

For the feasibility study, a software program was developed to manage and display data collected from the paver's on-board tag as well as to perform a number of intermediate computations.
Although this software has been run during the evaluation experiment described below, it is not yet refined to a point of being proposed as an industrial product. In the near future, the functions proposed by this software will also be partially provided by the OSYRIS on-board computer. Nonetheless, a dedicated system focusing on the tracability between plant and paving site has also been foreseen as a separate development.

5 INITIAL EXPERIMENT

5.1 Description

In November 2001, in cooperation with the DDE 77[1] offices, an initial experimental site was set up in order to demonstrate and validate this concept. For the experiment, we designed and built a prototype system around two distinct goals: quality documentation and operator assistance. This paper will focus primarily on operator assistance aspects, which may prove more illustrative and offer the advantage of introducing the system from a more attractive perspective for contractors.
The site consisted of a thin asphalt wearing course for a new road, located in the town of Bray-sur-Seine some 60 km southeast of Paris. A two-lane course of approximately 600 m x 6.2 m was built during the first day, and a traffic circle plus the adjoining access roads were built during the second day.

At the plant, due to the absence of a standard plug for connecting a quality-monitoring computer, we had to manually enter the data read on the plant control dashboard into a laptop computer and then transfer them into the electronic tag through the i-box.

All of the trucks involved in the experiment, five the first day and five the second, were equipped with an electronic tag screwed onto the right side panel, as seen in Figure 3.

[1] DDE 77 stands for "Direction Départementale de l'Equipement" (Regional Bureau of the Ministry of Construction and Public Works) for French Department No. 77 and is in charge of building and maintaining the department's public roads.

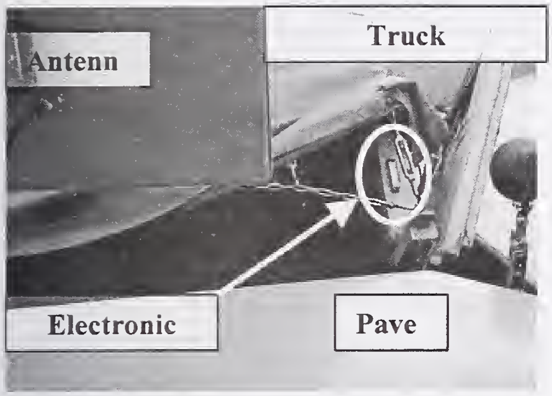


Figure 3: The tag screwed onto the truck side panel

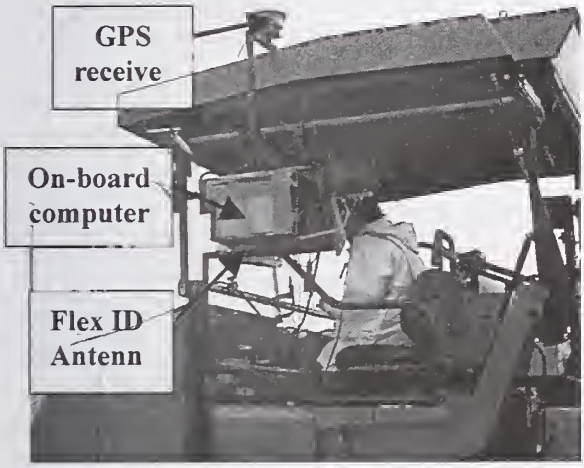


Figure 4: The paver on-board computer and GPS antenna

An i-box and antenna, placed in a suitable location for reading tags while the truck was downloading its material, were installed on the paver, as shown in Figure 4. In addition, an ultrasonic sensor was used both to detect the amount of material in the paver hopper and to validate completion of the downloading step. Figure 4 shows the on-board computer mounted on the side of the paver along with the OMNISTAR GPS receiver used.

5.2 Data recorded and displayed

The following data were selected to be recorded.

<u>Quality parameters, from the plant computers:</u>
- theoretical proportion of each component (in percent), - theoretical asphalt fabrication rate of the plant (in tons/hour), - bitumen temperature, upon introduction into the mixing tube (in °C), - asphalt temperature, when loaded into the trucks (in °C) (1), - precise bitumen weight for each batch (in kg).

<u>Quantity parameters, from the weighing station:</u>
- weight of the material batch for each truck (in tons) (2),(weight of the empty truck (in tons).

<u>Miscellaneous identification parameters:</u>
- identification of the work site, - identification of the type of mix design, - batch number, - identification number of the tag or truck (3), - starting and ending times of batch fabrication, - truck weighing time (4), - plane coordinates (using the national "Lambert" reference projection system) of the starting and ending points of the surface area covered by the batch, together with the application time, both provided by the GPS receiver.

In addition to parameters quoted (1) to (4), were displayed the following parameters, after computation by the on-board software:
- actual mass production rate of asphalt, using batch weight and fabrication time (5), - actual bitumen mass proportion for the material batch, using batch weight and bitumen weight (6), - average mass paving rate of the asphalt, computed from the batch weight and batch downloading time provided by the GPS receiver (7), - paving process output: ratio of paving rate to production rate (8), - exact average speed of the paver, thanks to GPS timing (9), - suggested correct speed of the paver, in order to increase output as much as possible (10), - average asphalt weight per unit area, using (9) and (10) and a theoretical width (11), - total cumulative asphalt weight (12), - total cumulative corresponding paving length (13), - estimated post-compaction thickness (14), using (13) and the theoretical specific mass of the compacted material (15).
Numbers 1-6 are related to fabrication, and 7-15 to the asphalt-laying operation.

5.3 Sample system output for tracability purposes

Figures 5 and 6 show examples of simple graphical outputs that may be of direct interest to the project manager or architect/engineer for checking the quality of the work.

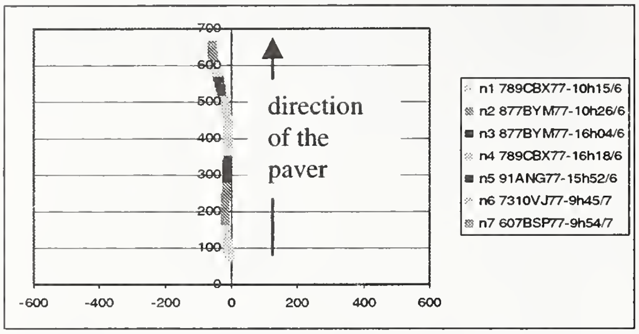


Figure 5: Location of the various batches along the road (straight line)

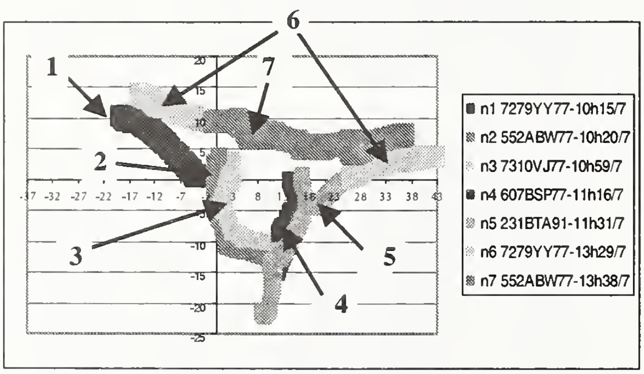


Figure 6: Location of the various batches along the road (roundabout)

Trucks are numbered in the chronological order. The scales used for both the *x*- and *y*-axes are in meters. This graphical display clearly provides valuable information on the way in which the site has been executed. For instance, it can be observed on Figure 5 that batch no. 5 was laid after batch no. 4, even though it had been loaded first at the plant. Given the availability of such output, we are convinced of site managers' and supervisors' ability to check very rapidly that site operations are running smoothly since any disturbance (e.g. abnormal delays between batches) can be detected quite easily. Figure 6 clearly illustrates the positioning accuracy obtained by GPS from traces recorded in the traffic circle, whose radius was about 10 m.

6 CONCLUSION AND OUTLOOK

This first site validated the satisfactory behavior of information transmission by means of RFID tags, which represented a novel technology for us, in contrast with the rather well-known GPS technology.

Despite the quality of results, this experiment was carried out in a semi-automatic mode and this created some problems at the site level by virtue of interfering with the execution and disturbing the process it is supposed to trace. It seems obvious that total process automation is necessary in order to attain the high level of reliability expected of the system.

In 2002, prototype improvements are foreseen to conduct further experimentation, in the aim of introducing a completely automatic mode of operations. These improvements will mainly concern software aspects. The new experimental site is scheduled to be operational by the end of the year: it will be a highway site, where we intend to trace all of the successive pavement layers.

7 REFERENCES

[1] Ligier, A. *et al.* (2001), Open System for Road Information Support, *Proc. Of the 18th International Symposium on Automation and Robotics for Construction*, Krakow, 205-210

[2] Finkenzeller, K. (1999), RFID Handbook - Radio-frequency identification fundamentals and applications, *John Wiley & Sons*, London.

Monitoring Construction Equipment for Automated Project Performance Control

by

Rafael Sacks[1,2], Ronie Navon[2], Aviad Shapira[2] and Irina Brodetsky[2]

ABSTRACT: The ability of company managers to respond to variations in performance in construction projects is severely limited by the time delay common in existing information and control systems in reporting schedule, budget and quality deviations from plan. We propose a system for interpreting data acquired automatically by monitoring the activity of major construction site equipment, such as tower cranes, concrete pumps, etc. The primary objective is to provide a reliable, real-time, automatic source of project performance information for construction managers. This requires interpretation of the monitoring results using a knowledge-based rule processor, and relies on the existence of a computerized Building Project Model.

KEYWORDS: Automation; Construction equipment; Information technology; Monitoring; Project performance control.

1. INTRODUCTION

The impact of management decisions on the work carried out on a construction site is severely limited by the time delay required in order to provide accurate reports of project performance in terms of schedule, budget and quality. In existing information and control systems, status reports are commonly delivered with very long time lags, resulting in undesirable trends being identified too late for corrective measures to be taken. Consequently, fully automatic monitoring of indirect performance parameters, and processing of the data to provide informative indicators of project progress, have been proposed (Navon and Goldschmidt 2001). Progress indicators can be compared to project control limits to deduce the project performance and identify problems. If the project information is directly available in the form of a computerized building project model, the interpretation can be done using knowledge-based software.

A focused research effort is currently under way at the Technion (Israel Institute of Technology) to develop the principles for:

- fully automated data collection on construction sites (Navon and Goldschmidt 2002),
- knowledge-based interpretation of the data in relation to a computerized Building Project Model (BPM) (Sacks 1998).
- provision of useful information about the schedule, cost and quality of the work being performed.

The generic term for systems using this approach is "Automated Project Performance Control" (APPC). The aim of such systems is to support the overall APPC goal of shortening the reporting cycle to one day, while eliminating the effort of reporting entirely. A number of projects have been initiated as part of this effort: a labor monitoring system using Automated Data Collection technology (Navon and Goldschmidt 2001, Navon and Goldschmidt 2002, Sacks et al. 2002), a system for monitoring and reporting work performed by heavy earthwork equipment (Navon, Goldschmidt and Shpatnitzky 2002), and a material control system.

Nearly all of the materials and components for a building are lifted into place by construction equipment, such as tower and mobile cranes, concrete pumps, hoists, etc. Tracking and recording the activity of such equipment is technologically straightforward. For these reasons, the authors propose that monitoring construction equipment holds the potential to provide significant data for APPC.

The system solution is a typical monitor and control loop (see Figure 1). The main part the control system is a decision rules processor. The processor receives data daily from the

[1] College of Architecture, Georgia Institute of Technology, Atlanta GA30332-0155 USA. Tel: 404-894-0437 rafael.sacks@arch.gatech.edu

[2] Construction Engineering and Management, Faculty of Civil Eng., Technion – Israel Institute of Technology, Haifa 32000 Israel Tel: 972-4-829-2600 ronie@techunix.technion.ac.il

monitoring 'black box' installed on the equipment. This input consists of load weights and crane-hook location coordinates gathered from the equipment through its working day. The location measurements are translated into the local site coordinate system so that they can be related to particular building elements or activities. The rules processor draws on a knowledge base of typical historical data, such as unloading durations or profiles for different construction activities, typical weights and heights of auxiliary equipment (between the hook and the load), spatial volumes typical of construction activities, etc. The data describing the project itself are drawn from a Building Project Model based interface, including details of pending activities (start/finish dates, duration, main material components and their approximate weights), probable storage areas and material delivery zones, etc.

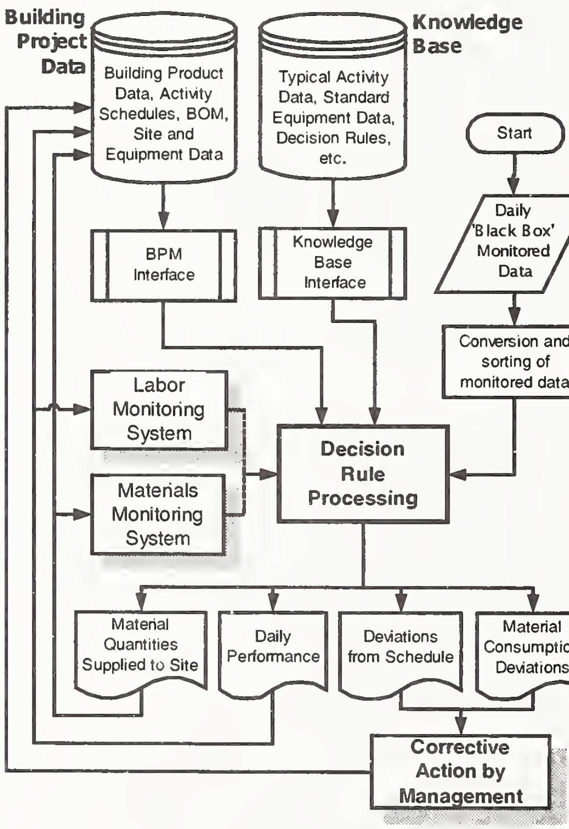


Figure 1. Construction Equipment Monitoring System

Using these three sources, the system is required to correlate the equipment movements with particular activities. It can then draw conclusions, with some calculated level of certainty, about the performance of any particular activity, and can provide additional information for other existing control systems (such as Materials and Labor Control Systems). Finally, the information provided is made available to management, and is used to update the project database.

In the following sections, the technology for monitoring equipment is presented for the typical case of construction tower cranes. Next, we detail the system information flow and the logic required for interpreting the data collected, and outline the implications for detailing of the necessary Building Product Model objects and relationships. Finally, we discuss the potential modes of use and benefits of equipment monitoring APPC systems.

2. AUTOMATED MONITORING OF CONSTRUCTION EQUIPMENT

The particular case of tower cranes is used here as an example (the methodology and technologies for other equipment types are similar). Selection of the monitoring equipment is dependent on the data required. These are identified as:

$W_g(t)$ – the gross weight of the load being lifted measured continuously through time.

$L_h(t)$ – the location of the hook in local building coordinates measured continuously over time.

Typical plots of $W_g(t)$ and $z_h(t)$ - the height coordinate of $L_h(t)$ - are shown in Figure 2. At least two distinct technological possibilities exist for measuring these:

<u>a) Separate Load and Location Measurement.</u>
A variety of tensile links, crane-scales, load measuring pins, load cells and others are available for measuring the gross load weight with maxima ranging from 500 kg to 100 tons (MSL 2002, Strainstall 2002). Most have cable or telemetry connections to a central interface block. Their accuracy is typically about ±1% of actual readout. A data recorder must be added. For location monitoring, a Global Positioning System (GPS) receiver may be used (Roberts et al. 1999). A pair of receivers may be used and operate in differential mode to avoid transmission interference, and to

achieve sub-metric accuracy (Peyret et al. 2000). GPS coordinates can be translated into local coordinates (Navon and Goldschmidt, 2002).

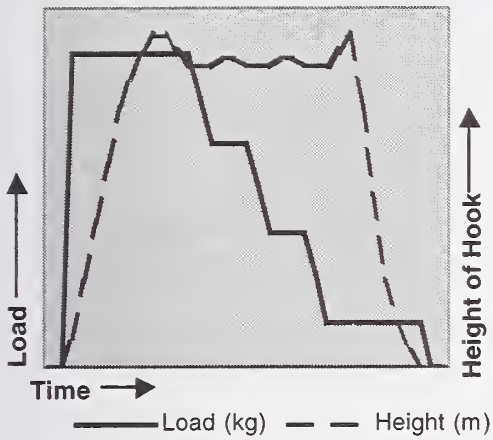


Figure 2. Typical Concrete Bucket Loading and Unloading Profile for Pouring a Slab.

b) Built-in monitors supplied by the equipment manufacturer. Large construction plant companies offer optional crane monitoring systems that provide real time monitoring of weight and location parameters. For example, Potain's '***Dialog Visu***' control system enables real-time monitoring and recording of the 'on-hook weight', and of the hook position in terms of boom rotation angle, radius of the trolley from the tower, and height of the hook from the ground (Potain 2002). (This definition of location, about the tower axis, can be considered a 'local crane polar coordinate system'). It should be noted that these advanced control systems are intended mainly to enhance site-safety, for more accurate crane operation in zones with reduced visibility, and for operation with remote control. They can be programmed to give warning against overload or to prevent travel into dangerous zones. They were not intended for project performance control purposes.

3. INTERPRETING MONITORED DATA

Comparing these data with calibrated typical load data and with the project model information should allow the system to draw conclusions about the type and identity of the building element being worked on the material or equipment being lifted, to identify supply of material to the site, and to identify the specific construction activity being performed.

3.1 Knowledge Based Rules

The knowledge rules encompass all of the processing as laid out in
Figure 3. It is envisaged that most of these will use production rule format expressions of the logic, although other techniques may be employed as needed. For example, consider derivation of the identity of the load, which is a central piece of information. Not only can the load itself be compared with known loads, but the pattern of unloading over time may also be matched with a library of typical patterns, such as the pattern for unloading concrete from a bucket for pouring a slab, as shown in Figure 2. In the figure the multiple points of unloading, and the residual load of the bucket during descent, are typical of casting concrete using a bucket on a crane. Recording a library of such patterns, and deriving pattern-matching rules from them, is therefore one of the tasks required for development of the knowledge base.

3.2 Project Data Model Information Requirements

Processing the knowledge-based rules outlined above requires that up-to-date information describing the construction project be available in computerized form. Given the diversity of sources of this information, we assume that it will be provided by means of a Building Product Model (Eastman 1999). The information must cover three main areas:

- the building **product:** the physical parts of the building.
- the construction **process:** the activities through which the building is built.
- the **resources** employed in construction (materials, equipment and labor).

The details of these are illustrated in the following discussion using the example of construction of a simple reinforced concrete column.

First, the existence of the column, in a particular geometric location in terms of the building's local coordinate system, and having a specific shape, must be known, so that the location of delivery of equipment and materials

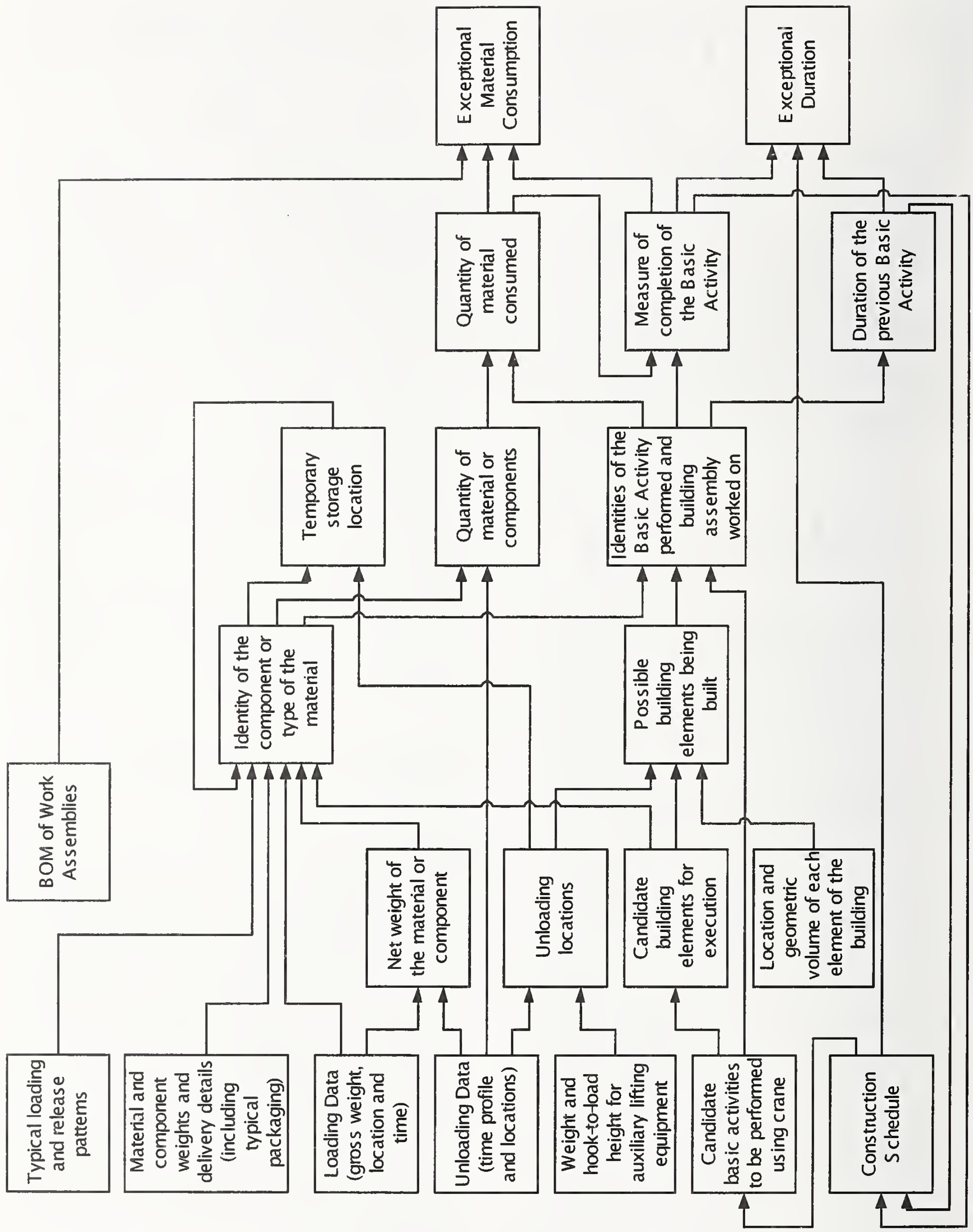


Figure 3. Rule processing flow.

can be correlated with a specific building element. The physical components of the column – the reinforcing cage and the concrete – must also be identified and described in sufficient detail to allow calculation of their weight and location. The fact that the column is part of an assembly consisting of the other RC columns on the building's floor (and possibly also the slab it supports) must also be known. These basic aspects of building product information – **product identity, location, geometric shape, material type** and **assembly/component aggregation** – are included in Building Product Models such as the IFC 2.x schema (IAI 2001). They are communicated in the CIS/2 structural steel product model (CIS/2 2001).

Second, the nature of the **construction method** and the planned **schedule** of construction activities must be available. At the simplest level, knowing whether the column is precast or cast-in-place is critical to interpreting the lifting data. Each construction activity, such as 'build 4th floor columns' must be fully defined in terms of its **scope** (the elements it includes - which columns are to be built), its **scheduled start** time and **duration**, the activities which **precede or succeed** it, and the **resources** that will be used and/or consumed in its execution (see next paragraph). However, this level of detail is insufficient to allow conclusions to be drawn using the rules outlined above. The **basic activities** of each activity (as defined by its construction method – Sacks and Warszawski 1997), must also be detailed (Sacks et al. 2001). For an assembly of cast-in-place RC columns, the activity 'Build Columns' would include the following basic activities: prepare column forms, fix rebar cages in place, position and stabilize forms, pour concrete, compact concrete and strip the forms. It is at this level that the APPC system contemplated here can identify execution of an activity – by associating the specific loads and unloading patterns of equipment and materials with specific sequences of basic activities. Some of these aspects of the construction process are provided in computerized construction schedules (note that the problem of updating schedules is specifically addressed by APPC. Nevertheless, while construction managers will be relieved of the need to input historic data, they will still be required to update their future intent).

Third, the resources of interest are those lifted by the crane – equipment (such as steel forms for RC columns) and materials (rebar cages and buckets of concrete). Their **identity**, **function** (i.e. association with a basic activity type), and **weight** must be detailed. In all cases except for bulk materials, their **shape** and **center of gravity** are required. Some resources, such as precast concrete pieces (elements), rebar cages, etc are project specific, while others, such as steel shutters, concrete buckets, etc. and all bulk materials, are generic across many projects.

Note that certain prefabricated parts of a building are defined as 'products'. For the proposed system to function, the fact that they also exist as material 'resources' (prior to their erection in the building) must be stored explicitly.

4. POTENTIAL BENEFITS

The processed information will be used for progress monitoring to produce 'as-made' information for updating project schedules. Additionally, it will be used as input to existing models – e.g. labor inputs control (Navon and Goldschmidt 2002) – for cross-reference purposes, thus improving their decision-making algorithms. Additionally, the research is expected to explore the potential for generating significant additional information from monitoring lifting equipment, such as:

- Accurate quantities (weights) of materials supplied (e.g. precise quantity of concrete poured), steel installed, etc.
- Patterns of building structure weights and their distributions (for long term validation of structural design codes).
- Potentially life-saving real-time information of crane hook location vis-a-vis worker locations (monitored with hard-hat GPS – Navon and Goldschmidt 2001).
- The identity and quantity of materials delivered to the site and the movement of materials within the site, for materials control purposes.
- Data regarding the availability of equipment and materials necessary for scheduled activities prior to their

commencement, to avoid delays and reduce downtime for crews.

- Data that can be used for improving the efficiency of equipment operating procedures, and/or of other aspects of the production process (e.g. work study data).

5. CONCLUSIONS

The concept of automated monitoring of construction equipment, within the framework of an Automated Project Performance Control system, holds the potential to greatly enhance the ability of construction managers to respond quickly to project performance problems. A road map for both the measurement and information technologies that are required has been set out. Ongoing research will focus principally on developing and testing the knowledge-base that must interpret raw monitoring data and produce reliable performance information.

6. REFERENCES

Ciesielski, C. A., (2000). "ADCIC - Automatic Data Collection in Construction." *A.I.M.'S Global Academic & Scientific Conference*. Cairo (Egypt).

CIS/2 - CIMsteel Integration Standards 2, (2001). http://www.cis2.org

Eastman, C.M., (1999). "Building product models: computer environments supporting design and construction," CRC Press, Boca Raton, Florida.

IAI (2001). International Alliance for Interoperability (http://www.iai.org), Industry Foundation Classes IFC Release 2.0.

Monitoring Services (UK) Ltd. (2002). "Load Transducers & Telemetry Monitoring Systems", www.msl.ision.co.uk

Navon, R., and Goldschmidt, E., (2002). "Automated Data Collection for Labor Inputs Control," *Automation in Construction* (accepted).

Navon, R., and Goldschmidt, E., (2002). "Can Labor Inputs be Measured and Controlled Automatically?" *ASCE Journal of Construction Engineering and Management* (accepted).

Navon, R., Goldschmidt, E., and Shpatnitzky, Y., (2002). "Field Test of an Automated Location Measurement for Project Performance Control," Special issue on Automated Data Capture Technologies of the *Journal of Computer-Aided Civil and Infrastructure Engineering* (accepted).

Peyret, F., Betaille, D., and Hintzy, G., (2000). "High-Precision Application of GPS in the Field of Real-Time Equipment Positioning. *Automation in Construction*, 9: 299-314.

Potain S.A. Ltd. (2002). "Lifting Equipment Products", www.potain.com

Roberts, G. W., Dodson, A. H., and Ashkenazi, V., (1999). "Global Positioning System Aided Autonomous Construction Plant Control and Guidance". *Automation in Construction*, 8: 589-595.

Sacks, R., (1998). "Issues in the Development and Implementation of a Project Model for an Automated Building System," *International Journal of Construction Information Technology*, Salford University, Salford UK Vol. 5 No 2 pp. 75-101.

Sacks R., Navon, R., and Goldschmidt, E., (2001). 'Building Project Model Support for Generating Automated Project Performance Control Information', submitted to the *ASCE Journal of Computing in Civil Engineering*, December 2001.

Sacks, R., and Warszawski, A., (1997). "A Project Model for an Automated Building System: Design and Planning Phases," *Automation in Construction*, Elsevier Science Publishers B.V., Vol. 7 December 1997 pp. 21-34.

Strainstall (UK) Ltd. (2002). "Industrial Load Measurement", www.strainstall.com

Experiences with the Design and Production of an Industrial, Flexible and Demountable (IFD) Building System

Frans van Gassel *

Eindhoven University of Technology, Department of Architecture, Building and Planning, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

ABSTRACT: The Dutch government encourages innovative construction by subsidizing cohesive industrial, flexible and demountable building (IFD) pilot projects. Industrial building concerns the process-related aspects of production, robotization, mechanization, automation, prefabrication, communication, etc. Flexible building involves products that are made in accordance with customer's wishes and the possibility to make adjustments when the building is in use. Finally, demountable refers to the sustainability of the building.

This paper describes the experiences relating to the design and production of a demonstration project (IFD Today) for an apartment building system on the Eindhoven University of Technology campus. A housing corporation had commissioned the design of an industrial, flexible and demountable building system for flats, on an existing foundation, the surface area and floor plan of which had to be adaptable during use. The current flats do not meet present requirements as set by the government and the occupants, and renovation is too expensive.
The data was obtained through participation in the design team and close production monitoring. The paper focuses on industrial building.
After monitoring the design and production process, the following conclusions can be drawn:

- IFD building requires co-operation and a multidisciplinary approach during the design process. Essential issues to be considered during the design process are the tasks, choice of designer, design tools and the expected result. Design meetings must also be organized.
- Due to the lack of a suitable model to calculate the production, operating, renovation and demounting costs and the absence of a marketing plan, conventional designing solutions were frequently applied.
- The steel structure and floor panels were mounted very quickly. This was in contrast to the outer walls, roof, fittings and finishing elements.
- It is advisable to entrust one specific company with the responsibility for a number of production tasks, such as planning, making and checking the drawings, and allocating labor resources with regard to transport and safety.

KEYWORDS: Building Site, Building System, Industrial Construction , Multidisciplinary Design.

1. INTRODUCTION

A trial module was built on the Eindhoven University of Technology campus to gain experience with the design and production of an IFD building system. This paper presents the findings. The first paragraphs explain the concept of IFD building, the IFD Today project and how the study was set up. Finally, conclusions and recommendations are made with regard to the marketing of the IFD building system, based on the results of the study.

2. IFD BUILDING

In the Netherlands, a program entitled 'Industrial, Flexible and Demountable Construction Demonstration Projects' has been developed by the Ministry of Economic Affairs and the Ministry of Housing, Spatial Planning and the Environment. This program was developed to promote the innovative application of industrially developed and manufactured construction components in the construction and renovation of homes and public utility buildings.

* E-mail: f.j.m.v.gassel@bwk.tue.nl

IFD building provides an integrated approach to the initiation phase, the design, production and use of buildings, and is characterized by early co-operation between the parties, enabling alignment of the concept, design and execution. The industrial production of components offers increasing opportunities for flexible use. Demountable building enables a separate replacement of components with various lifespans, thereby extending the life of the building as a whole. As such, IFD building is a form of sustainable building. [SEV]
Van den Brand describes IFD building as a three-pronged strategy to innovate the building process: the client (flexible), the manufacturer (industrial) and society (demountable). This is characterized by the following principles: 'level-thinking' [building - floor - room - work station], a fixed form and space combined with variable options for the interior, a multidisciplinary design, separate technical systems; dimensioning and nodes; modularity and demountability. [Van den Brand]

3. IFD Today

Between 1945 and 1963, about 400,000 apartment buildings were constructed in the Netherlands with the following characteristics: traditional design, the use of clay, no insulation, poor sound insulation. They were, almost without exception, four-storied staircase-access flats. Between 1964 and 1975, another 475,000 medium-rise apartment buildings and 300,000 high-rise apartment buildings were constructed. Today, these apartments still have poor (sound) insulation (see figure 1).

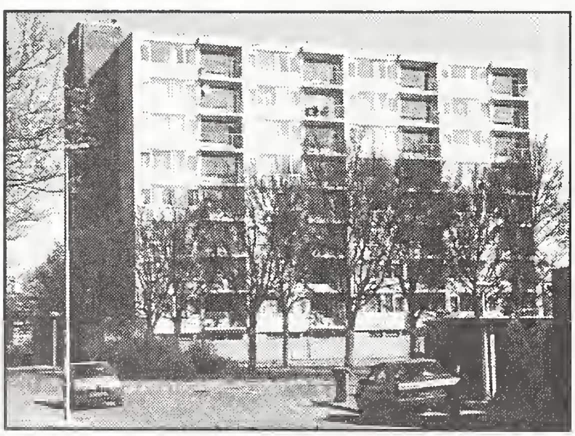


Figure 1. Apartment blocks.

The improvement of post-war houses as from 1974 mainly involved the application of insulation. Although heat insulation was improved, the number of problems associated with building physics increased. In spite of substantial investments, very little could be done to improve the living quality. Houses were cramped and inadequately equipped, they had poor sound insulation, and were only accessible by a narrow staircase.
A more drastic approach proved technically impossible or would have required substantial and unprofitable investments. The current demand for houses is very diverse in terms of surface, layout and fittings. Even in cases where a drastic approach was adopted, the required level of housing differentiation could not be achieved within the existing shells. For this reason, more and more post-war houses will be demolished. [IFD Today]
A partnership (IFD Today) between the Amnis housing corporation, contractor Heijmans IBC, installer Stork and the Eindhoven University of Technology has chosen to solve the above problem with IFD building. A design team comprising staff employed by the partners was instructed to develop an IFD building system using existing technology.

4. IFD BUILDING SYSTEM

According to Eekhout, a building system is an orderly collection of construction elements and construction components with connecting facilities, which can be combined or applied in various ways, in accordance with regulations or agreements and depending on the environment. [Eekhout]
The building system process is characterized by the following aspects: technology, the human factor, information and organization. The building system product is characterized by the aspects of function, flexibility, geometry, materials, structural and technical capacity, complexity and cost. The dominance of the aspects depends on the market in question. The production of frequently used construction elements has now largely been mechanized or is carried out by robots, for example. Specific construction elements requested by the client are manufactured in the traditional way.

The conceptual design in IFD building is based on six specific values of a building:

- Basic value: achieving a building physics level that is higher than current standards.
- Use value: aiming for as much layout freedom as possible at building, story and apartment

level and with regard to individually adjustable and quantifiable installations.

- Local value: achieving a building system that allows differentiation possibilities according to type, surface, layout, fittings and architecture.
- Ecological value: reusing foundations and the necessity of applying light structures.
- Economic value: developing a product in line with marketing conditions.
- Strategic values: constructing a building that can be adjusted in the course of time.

These values can be considered on four different levels: the built environment, building, house and living quarters level. The conceptual design addresses the design aspects, viz. the requirements, design, quantification and design strategy.
[Rutten] [Systeemcatalogus IFD Today]

Hendriks has developed a number of design criteria for IFD building:

- Integration and independence of disciplines: installation, bearing structure, outer shell and interior finishing.
- A completely dry building method: no pouring of concrete, mortar joints, screeds, stuccowork, sealant or PUR spray.
- Perfect modular dimensioning: a great deal of attention to drawings, prototype testing, quality system for drawings, and assembly instructions.
- Adjustability of all parts: bearing structure (limited), installation (practically unlimited), outer shell (limited and modular), interior finishing (practically unlimited and modular). [Hendriks 1999]

The IFD building system has the following characteristics:

- A steel support construction of hot-rolled standard profiles, L-supports at the corners, and I-supports and joists on the grid line.
- The supports provide stability and the joists have been fixed to the bearers in such a way that vibration is kept to a minimum. The floor panels are placed on the joists.
- The floor panels are hollow to enable the inclusion of technical installations. See figure 2.
- Piping can be placed in the supports and the floor panels at the plant or the building site.
- Non-bearing walls dividing houses and rooms can be placed anywhere.
- Free choice of outer wall.
- There is no specific roof design. The traditional solution is adopted, i.e. warm roof.

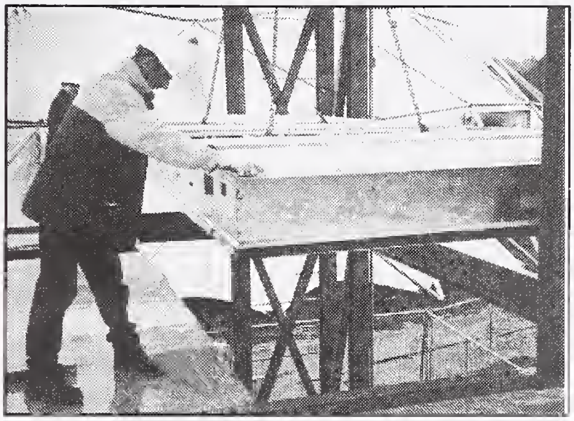


Figure 2. Assembling the floor panels.

Characteristic elements of IFD building are:

- An 11-meter span and a 7.2-meter grid.
- Optimal free space providing various layout possibilities.
- Maximum flexibility with respect to vertical and horizontal piping, providing various possible locations for toilets, kitchens and bathrooms.

[Systeemcatalogus IFD Today]

5. STUDY

To gain experience with the IFD building system, a trial module was built on the DUBO (Sustainable Building) park on the Eindhoven University of Technology campus. The trial module is fourteen meters wide and eleven meters deep, and has two stories. See figure 3.

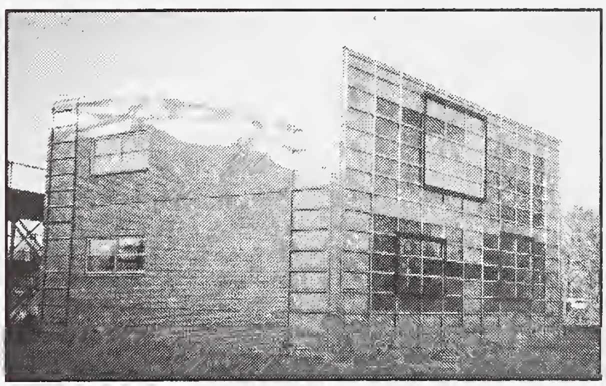


Figure 3. IFD building system trial module on the Eindhoven University of Technology campus.

Not only the trial module's design process but also its production process was assessed for its IFD characteristics.

This report only describes one aspect of IFD building, viz. industrial building.
According to its conceivers, the IFD technology makes use of the following traditional elements of industrial building:

- Production at the plant.
- Mechanized production.
- Mass production.
- Co-operation is independent of projects.
- Deployment of information technology.

Today, the emphasis in industrial building is on:

- Customer-oriented production and marketing.
- Flexible production systems.
- Subsystems are independent but can be combined.

[Hermans]
The transfer of production from the building site to the plant has resulted in an integration of functions and sometimes materials of a component or a system of components. This requires a high level of component alignment (specifications, dimensioning, finishing) [Bouwen op kennis]
For the purposes of this study, the following description of industrial building is used: the transfer of physical tasks from the building site to the plant and from people to machines and computers. These tasks do not involve building alone but include the collection and processing of data, consultations and alignment, optimization and planning.

It is apparent from this description of industrial building that designers from various disciplines – architectural designers, building physics designers, electrical and mechanical designers, construction designers and implementation designers – are involved in joint designing from an early stage.

5.1 Design process study

The trial module was designed in close co-operation by a team consisting of an employer, architect, structural engineer, fitter, contractor and representatives of the supply industry. An attempt has been made to describe the experiences gained through intensive participation in design meetings:

- The preparation for the meetings was limited to an agenda.
- Other subjects, such as planning, subsidy applications, composition of the design team, were also discussed.
- The fitter, in the capacity of installation consultant, aims to find quick solutions.
- The fitter discusses solutions a great deal.
- Little use was made of sketches in the concept design phase.
- Design problems were passed on to the subcontractors, who then solved them independently.
- Production costs were calculated in the traditional manner.
- Design problems were still being solved in the execution phase, usually at the building site.

During the building phase of the trial module, the meetings between the designers and the parties executing the work were sometimes reminiscent of a site meeting.
During the design and production process, a list of points for improvement was made, which could be used during a subsequent design process.

5.2 Production process study

The production process is considered a transformation of materials into a building, with waste as a by-product. The transformation is carried out by people and tools and managed using plans, drawings and instructions (monitoring). See figure 4.

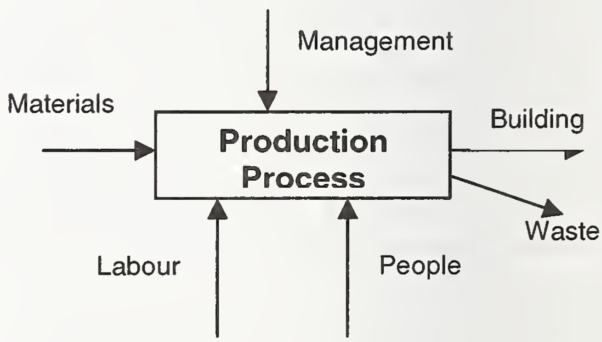


Figure 4. Model of the production process of a building.

The following aspects of the production process can be observed:

- The production process: preparation, transport, storage, treatment, processing, conditioning.
- Materials, construction elements, construction components.
- The building and (separate) waste flows.
- The people: planners, project manager, supervisors, construction workers, mechanics.

- Labor resources: tools, scaffolding and accessibility structures, transport equipment.
- Management: planning, schedules, instructions, drawings.

The findings obtained during close monitoring of the production process at the plant and at the building site were recorded in a log, filmed or photographed. Explanations, comments and points for improvement are provided for photographs depicting extraordinary situations. Table 1 gives a concise overview of the findings and points for improvement for each aspect. [Van Gassel]

6. CONCLUSIONS

The following conclusions can be drawn after monitoring the design process and production process:

- In IFD building, the design process requires co-operation and a multidisciplinary approach. Matters such as design tasks, choice of designers, design tools and expected results must be considered during the course of the design process, and design meetings must be organized.
- Due to the lack of a suitable model to calculate the production, operating, renovation and demounting costs and the absence of a marketing plan, conventional design solutions are frequently chosen.
- The steel structure and floor panels are demounted very quickly. This is in contrast to the outer walls, roof, fittings and finishing elements.
- It is advisable to entrust one specific company with the responsibility for a number of production tasks, such as planning, making and checking the drawings, and allocating labor resources for transport and safety.

7. RECOMMANDATIONS

The IFD building system should be treated as a comprehensive product and be marketed by one company or a cluster of companies. The company or cluster of companies should develop an IFD cost model and marketing plan for this purpose. The marketing plan should address at least the following elements: living concepts, mass customization, adaptability at all phases of the building's economic life, period of use versus economic life, customer order release point, long-range production planning etc. Specific skills required for the organization and facilitation of and participation in design meetings could be obtained by a mutual exchange of knowledge.

8. REFERENCES

[Van den Brand] Van den Brand, G., Rutten, P., Dekker, K. 1999, *IFD Bouwen, Principes en Uitwerkingen.* TNO (Netherlands Organization for Applied Scientific Research) report 1999-BKR-R021, TNO Bouw (Building and Construction Research) Delft.

[Bouwen op kennis] *Bouwen op kennis, Rapportage verkenningscommissie Bouw*, 2000, Adviesraad voor het Wetenschaps- en Technologiebeleid.

[Eekhout] Eekhout, Mick, 1997, *POPO of ontwerpmethoden voor bouwproducten en bouwcomponenten*, Delft University Press.

[Van Gassel] Van Gassel, Frans, 2001, *Bijdragen aan de IFD Today systeemcatalogus.* Report of the Eindhoven University of Technology.

[Hermans] Hermans, M, 1997, *De marktpositie van IFD-bouwen voor de Nederlandse Bouwindustrie*, Damen, Rotterdam.

[Hendriks 2001] Hendriks, Nico, Frans van Gassel, 2001, *Construction of a prototype of an industrial, flexible and demountable (IDF) apartment building system. International Status Report on Aspects of FutureSite.* CIB Publication 265, Wellington NZ.

[Hendriks 1999] Hendriks, Nico, 1999, *Industrieel, Flexibel en Demontabel Bouwen (IFD): Ontwerpen op veranderbaarheid*, Eindhoven University of Technology.

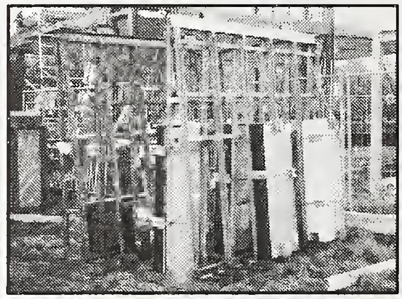
[IFD] IFD Today, 2000, Leaflet.

[Rutten] Rutten, P., Trum, H., 1998, *Prestatiegericht ontwerpen en evalueren.* 'Geïntegreerd ontwerp van gebouw en installaties', reader of the Eindhoven University of Technology.

[SEV] SEV, Leaflet IFD and www.sev.nl

[Systeemcatalogus IFD Today] Internal IFD Today publication.

Table 1. Findings and points for improvement with regard to the production process.

Aspects	Findings	Points for improvement
Materials	• Additional production space was needed at the plant to mount technical installations into the steel construction. • Mechanics had to determine reference points and dimensions themselves as these were not indicated on the drawings submitted. • The production of large elements was treated as a project instead of serial production. • Mechanics frequently tried to mount installations at the building site instead of at the plant. • Parts were selected and the assembly order was determined at the building site. • Small materials were delivered unsorted and in bulk. As a result, much sorting was needed and materials got damaged.	• Special assembly drawings should be made. • Investments should be made in mechanization and robotization. • Construction components should be coded. • Materials should be supplied in crates and using additional means of transport. 
Management	• Each supplier submits its own drawings. Connections between construction components of the various suppliers were not provided and a solution had to be found on site. Many telephone calls were needed to obtain the required assembly information.	• Drawings should be made of the connections between the construction components. • Specifications should be of a quality that the need for drawings and instructions is eliminated. • A project database and project extranet should be used.
People	• Different types of scaffolding were erected for different parts of the work. • Traditional safety provisions such as edge protection were used.	• One type of scaffolding should be used. • Safety provisions should be integrated in the building system.
Labor resources	• Various types of cranes were used, sometimes as many as four at the same time. • Special labor resources were made and deployed at the building site for the purpose of the project.	• One type of crane should be used during the execution process. • Specific labor resources should be designed.
Waste	• Much waste was produced during the fitting-out process. • Waste was not collected separately. 	• Returnable packaging should be used. • Work on material carried out at the building site should be kept to a minimum. This should be confined to the plant as much as possible. • Waste should be separated in accordance with government regulations.
Building	• Building physics measurements showed that a number of details had not been finished properly: noise leaks, air leaks, etc.	• Details should be finished properly.

FutureHome – A PROTOTYPE FOR FACTORY HOUSING

by

Robert Wing[1] and Brian Atkin[2]

ABSTRACT: The manufactured housing project ***FutureHome***, which received major funding from the European Commission, has developed the engineering know-how to create affordable, high quality, cost effective manufactured housing, with a customer focus that takes account of diversity of styles, designs, materials and locations. The provision of housing to an acceptable standard is becoming an increasingly serious problem worldwide, and the widening gap between supply and demand points to manufactured solutions as the most viable way forward.

FutureHome has developed systems for product and process analysis suited to manufactured and prefabricated construction solutions, and these have led to design prototypes. To mark the end of the research phase of the project, one design has been built to full scale as a demonstrator. The research has led to leaner design and construction processes that focus on value for money, improved productivity, maintainability and sustainability. It has also focused on basic engineering requirements such as fast connectors, and automation solutions for materials handling and assembly. Further aspects of the research programme have considered how the product can be adapted to areas of high seismic risk, and has developed software solutions for design and management from a virtual-reality customer interface to cyber-agents for logistics.

FutureHome is expected to have benefits through savings in construction costs and time, significant reductions in defects on completion, and will enable industry to be more competitive overall. Other benefits include improving the quality of life, social fabric and health of the European economy through a more efficient and effective construction process.

Keywords: housing; manufacture; construction; automation; process

1. INTRODUCTION

FutureHome is a European R&D project bringing together fifteen partners from six European countries, and forms part of a global project under the Intelligent Manufacturing Systems (IMS) programme. This project has explored the potential to apply advanced manufacturing technology to housing design and production, aiming for significant construction cost and time savings, and major reductions in defects on completion. The studies have shown that high quality can be delivered at a fraction of the cost of the traditional construction methods that currently dominate the market. A strongly IT oriented approach is seen to be fundamental, with state of the art simulation tools for support of the design and construction processes and effective involvement of the customer.

The project is industry-driven and draws on the substantial expertise and resources of some of Europe's major manufacturers and niche contractors – all with a reputation for developing and applying innovative solutions.

The European Commission supports the project under its *Brite*-Euram programme, enabling leading European research institutes and universities in this field to work alongside the industrial partners; the European consortium has a formal collaboration agreement with IMS project IF7, "The Intelligent Field Factory", involving a large Japanese research consortium. The Japanese project is primarily concerned with larger,

[1] *Department of Civil and Environmental Engineering, Imperial College of Science, Technology & Medicine, London SW7 2BU, United Kingdom. Tel: +44 20 7594 5997 Email: r.wing@ic.ac.uk*

[2] *Atkin Research & Development Ltd., Woodcote, Reading RG8 0QX, United Kingdom Tel: +44 1491 681818 Email: brian@atkin.co.uk*

particularly taller, structures, but in other respects its objectives parallel those of ***FutureHome***. That consortium includes three major construction companies in alliance with universities and research institutes *[Ref. 1]*.

2. HOUSEBUILDING AND EUROPE'S ECONOMY

At the start of the 21st Century, much of manufacturing industry is exploiting the efficiencies of mass-customisation, with design integrated to computer-controlled production. Customers enjoy affordable high-tech goods and, particularly in the fields of communications and entertainment, see improvements to their quality of life and work. The housebuilding industry, however, is slow to adopt the benefits of our post-industrial digital age, opting for the traditional craft approach with its rather variable output quality.

Attempts during the 20th Century have failed to provide a factory approach to housing construction, and today's most advanced offerings, represented by pre-fabricated modular systems are struggling to gain their market. Factory-produced panel constructed housing is gaining popularity in the Netherlands and Germany (*Fertighaus*), where this approach has become a substantial industry sector *[Ref. 2]*. The main advantage is that many of the construction processes are brought indoors, and benefits then arise from reduced weather impact on scheduling, and the relatively easy availability of machines and services in a well-structured factory environment.

Extension of this principle into 3 dimensions is seen in modular prefabricated housing. Here entire room sized units can be completely fitted out within the factory and delivered to site for assembly. A demonstration project of this type has been recently completed in London (Murray Grove), where it was the client, Peabody, the UK's largest housing trust, that drove the initiative to use modular methods; the general success convinces it to use the same concept on further projects.

Such projects, however, represent only a minute proportion of the construction industry, the largest industrial sector in Europe's economy; with an output of 780 million Euro it exceeds Japan's construction industry by 10%, and that of the US by 30%. It is also Europe's largest employer, providing jobs for 11 million workers, and with each construction job generating another two in related sectors some 21% of Europe's workforce therefore depends directly or indirectly on construction.

In this time of sociological, demographic, and rapid technical change, the industry is adapting to many new challenges, including:

- *Creation of multi-functional buildings capable of adapting to the changing nature of work and leisure.*
- *Provision of housing for Europe's ageing population, offering comfortable and autonomous lifestyles.*
- *Provision of quality urban environments.*
- *Reduction of energy usage, buildings currently being responsible for more than 40% of total consumption.*
- *Reduction of waste; future building products need to be re-usable and recyclable.*

The above factors, together with client demands for both quality and quantity in the housing sector indicate that factory methods will be essential in the future if the industry is to handle future demand and avoid loss of this market to foreign competition. The greatest challenge for the industry is to leave traditional methods behind, and evolve solutions based on a fully manufactured house rather than simply introduce factory-prefabricated parts into current processes. The perceived solution is mass-customisation rather than custom-build, together with an IT-based approach to design and construction.

3. THE *FutureHome* CONCEPT

FutureHome's priority has been to evolve a design system that takes account of the diversity of styles, designs and material composition, and especially the preferences of owners and occupiers, that is, customers who are as diverse as Europe itself. All share the same basic requirement of a home that is affordable, decent, modern in its facilities and capable of adapting to changing needs. The general maintainability of the products of

FutureHome, employing materials and components that can be reused, is a key objective.

New approaches to design and production are needed, for assembly in a controlled, factory environment by a higher skilled labour force, using materials more efficiently, and involving low energy use in operation. These, in any case, are essential elements in Europe's path towards sustainable development.

Production has to be divided between fixed plants for component manufacture, and field factories for assembly of larger units local to the construction sites. The field factories reduce the overall transportation resources required, and importantly will bring employment to the area undergoing regeneration. It is not, however, an easy change for the traditional housebuilding industry, and it is possible that the initiative for setting up the manufacturing facilities will come from another industry sector, as has happened in Japan.

The current panel-based pre-fabricated housing systems use production processes based on labour rather than machines. They address part of the problem, as site-based employment becomes less attractive to prospective construction workers and consumers demand more consistent quality. Systems based on steel, timber and concrete structural components have been developed, but only steel and concrete would be suitable for the full range of apartment types and heights required, and the economical transport radius for concrete systems is limited. Thus the core components in a new European housing industry would need to be fabricated from steel *[Ref. 3]*.

Small-scale suppliers of pre-fabricated buildings based on light steel frames exist in most European countries. The challenge is to create a more efficient integrated production system, taking advantage of these firms' knowledge of local requirements as well as drawing on the expertise in other industrial sectors with long experience in combining customisation with mass manufacturing technique. Customisation is the key to the success of this approach; current housebuilding supply chains are not robust enough to tolerate the variabilities in customer demands because capacity is not engineered into the system.

The IMS link with Japan has been valuable in understanding the market requirements, as in that country mass-produced, customised homes have secured 7% of the new housing market, and with excellent customer response

Development of the new housing systems and associated manufacturing plant proposed in ***FutureHome's*** outputs will contribute to the regeneration of urban areas, create new skills, and establish integration of the technological, social, economic and business aspects to support the future provision of housing in Europe.

4. *FutureHome's* ACHEIVEMENTS

The engineering R&D within the project covered three linked areas:

- **Building System studies.**
- **Automation studies**
- **IT studies**

Examples of ***FutureHome*** innovations and developments from each area are given below; it is hardly necessary to state, however, that even this large project barely scratches the surface of the major gains obtainable from a full factory-based solution.

4.1 The *FutureHome* building system

The ***FutureHome*** building system derives from a Kit-of-Parts approach, wherein standardisation plays a significant role in achieving economies from the many variations allowed in the parts set. The choice of light steel framing for the pre-fabrication base comes from transport logistics and the requirement to use automation tools both in the factory and on site. This demands assured dimensional precision and avoidance of problems such as warpage and shrinking.

Kit-of-Parts is a specific implementation of prefabrication; the distinction being that such structures follow an assembly, disassembly, parts replacement, re-assembly sequence as required during their life-cycle. Normal prefabricated structures, however, can only be taken apart with loss of functionality, either due to unavoidable damage or irreversible jointing.

The general modular concept, I-KOP, or Integrated Kit of Parts, centralises the provision of services in standardised core modules, leaving considerable design freedom for the outer (less serviced) modules and panels *(Figure 1)*.

4.2 Production optimisation for prefabricated buildings

The balance between site and factory processes, and the optimum level of prefabrication for housing designs are analysed using software tools such as DSM (Dependency Structure Analysis), a systems analysis tool for the investigation of interactions and interdependencies between elements in a complex system – modules and sub-systems in this case. Used on the architecture of the product, DSM determines possible integrative components, in other words, it seeks potential for employing larger factory pre-fabricated modules. Used on the assembly process DSM creates optimised task sequences that can be fed into planning tools to determine the critical path for the assembly process.

Additionally, IDEF0 process analysis methods have been used to produce optimised workflow procedures for the purpose of providing a model of the total process for design and manufacture of ***FutureHome*** products. The aim has been to create a single, integrated model of the process from which views of the various phases can be mapped in detail. Information has been sought from the project partners and from outside the project to develop this blueprint for total process control.

The use of the international standard, IDEF0, to portray the logic of the process, its activities, information flows, controls (i.e. constraints) and mechanisms (i.e. personnel and tools) within the model brings together in one place the core information needed to organise and manage the design and manufacture process. *Figure 2* illustrates the top-level view of the process, which is decomposed into lower levels of detail that equate to specific work flows.

Using *BPWin* software, a detailed model has been created, with activity, cost and time data incorporated. The model has been used to export these data to an ODBC-compliant database (*Microsoft Access*), located on a password protected web site: the term Process Control Interchange has been given to this development. Authorised users can interact with the model's database to import data for cost estimating and time scheduling. Other data can be exported to assist in defining work flow and decision control procedures.

One of the benefits of this modelling approach is that a change in any one element is reflected automatically in the rest of the model, thus preserving data integrity. Different arrangements within the process can be tested and optimised.

The model combines an earlier model for the design of a structure at full size, with those for off-site manufacture and on-site production of the components and assemblies. This holistic view of the process provides the basis for the development of ICT tools for automating work flow. Output from the model can be interchanged with many applications, allowing value to be added to the basic process model.

4.3 Connectors for automated assembly

Complementing the Kit-of-Parts, connectors have been developed for ***FutureHome*** modules, covering structural, assembly, and services requirements. These are seen as the essential means of enabling the use of intelligent machines for assembly processes, and they make use of *Design for Assembly* and *Smart* concepts *[Ref. 4]*.

For site assembly of prefabricated modules using conventional construction site plant equipment, conical paired assembly connectors allow auto-centering of modules. A computer-controlled 1/5 scale crane was built for evaluation of this approach to site assembly of large modules *[Ref. 5]*. Snap-fit structural connectors for modules and panels have also been developed, using plastic compression spring fastenings.

A universal services connector provides a highly standardised factory fitted system, where traditionally a variety of tradesmen are involved on site. The object here was to develop a connection system that not only reduces pure on-site assembly time by use of a fast mechanical connection, but also brings a substantial reduction of the total assembly time by off-site integration of trades (*Figure 3*). It will thus be possible to increase added value in the factory and deliver components of large manufacturing depth and a high degree of prefabrication to the construction site. In addition to the cost advantages, quality improvement is assured.

4.4 Cyber-agent technologies

Cyber-agents as decision-making aids have many potential applications in both on and off-site construction processes. The project has investigated some specific examples, addressing agent communications issues extensively, covering agent-agent as well as agent-human operator interactions. The lack of international standards, security, and network bandwidth are seen as the main problem areas in their implementation.

A materials delivery agent was developed as a demonstration of this technology. Using electronic tagging devices attached to individual components, the delivery cyber-agent tracks parts, reconciles items with orders, and interfaces with the construction database. Information is stored in attached 'e-tags', which for the type selected have up to 64Kb memory.

4.5 Visualisation developments

Visualisation methods provide support on several levels in this project. VR is employed to complement the Kit-of-Parts, providing a common virtual environment to be shared by clients, architects, engineers, constructors, maintenance providers, etc. Thus clients can be led through the virtual design of their house, which is created on-line using prefabricated components from the Kit-of-Parts.

In addition to the house model, a simulation of the construction process of the house is provided, allowing examination of cost, quality and time parameters *[Ref. 6]*.

The design environment within which the user can create a design from its constituent prefabricated components *(Figure 4)*. A 3D window (top left) lists the components available from the current library. The user can select a component and view it in 3D; they can then choose to add the component to the design, using collision detection and constraint based modelling techniques to ease the interaction process. As the design process proceeds, a window within the environment keeps the user updated with the total component cost of the design.

The user then moves to the construction environment for an animated simulation of the construction process *(Figure 5);* this includes a tower crane and other virtual site equipment.

4.6 Performance and economics

Space limitations preclude adequate presentation of many of the innovative contributions from this large partnership, especially those relating to the use of IT. The research has included developments required to ensure a satisfactory whole-life performance for the products, and to deliver good value for money. Building physics issues have included structural analysis, thermal characteristics and ventilation issues, acoustics, seismic performance, regulatory, sustainability and ecological requirements – especially those regarding material choice and reuse. These, together with architectural, aesthetic, and environmental considerations are clearly essential to any design for future housing.

As the main phase of this project ends, the partnership believes that the ***FutureHome*** project has significantly progressed the arrival in Europe of a fully factory-based approach to housing that exploits the benefits of today's digital technologies. The use of visualisation and mass-customisation concepts will, we trust, avoid the greatest perceived danger, that of factory housing being associated with styles that are bland and incapable of expressing custom or culture.

REFERENCES

[1] Yagi J., Automation & Robotics in construction in Japan, Proc. 16th ISARC, Madrid, Sep 1999, p.III
[2] http://www.europahaus.de
[3] Davies J.M., Steel framed house construction, The Structural Engineer, Vol 78 No. 6, Mar 2000.

[4] Wing R.D., Factory Housing – Making the Connections. Proc. 17th ISARC,Taiwan, Sep 2000, Keynote, p. 33.
[5] Balaguer C. *et al*, FutureHome: An Integrated Construction Automation Approach, IEEE Robotics & Automation, Vol.9, No.1, Mar 2002, p. 55.
[6] Murray N. *et al*, A Virtual Environment for Building Construction, 17th ISARC, Taiwan, Sep 2000, p. 1137.

FIGURES

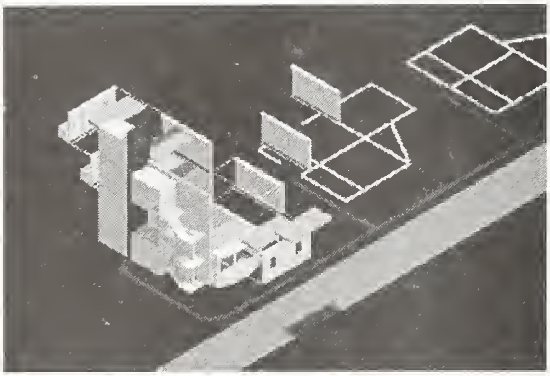


Figure 1. I-KOP Building system

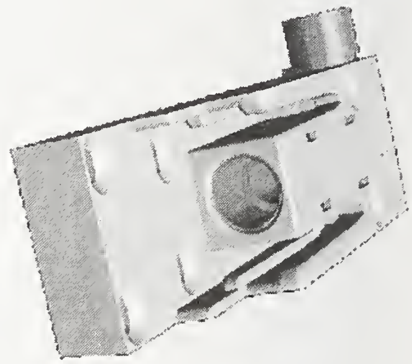


Figure 3. Services connector prototype

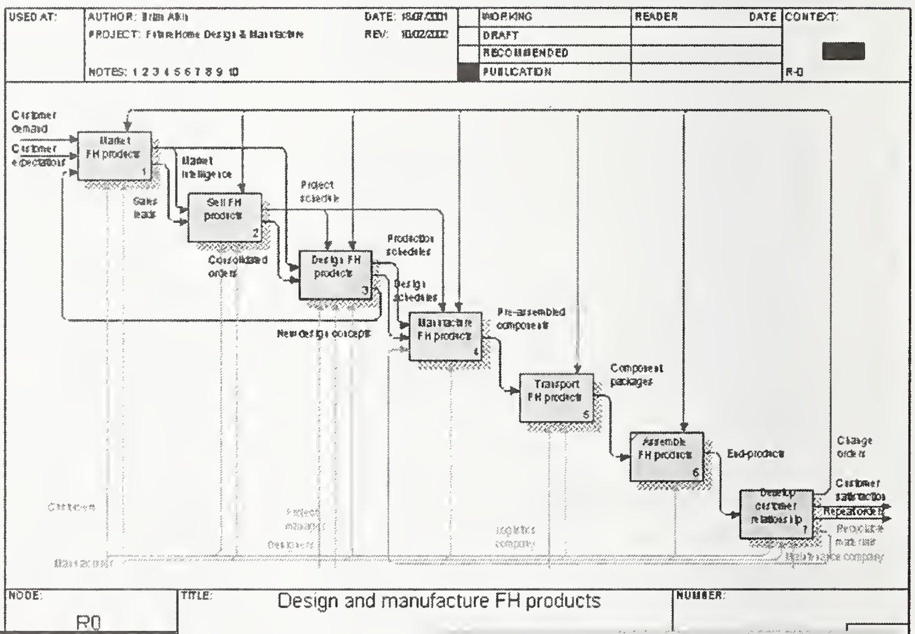


Figure 2.
Top level IDEF0 process diagram

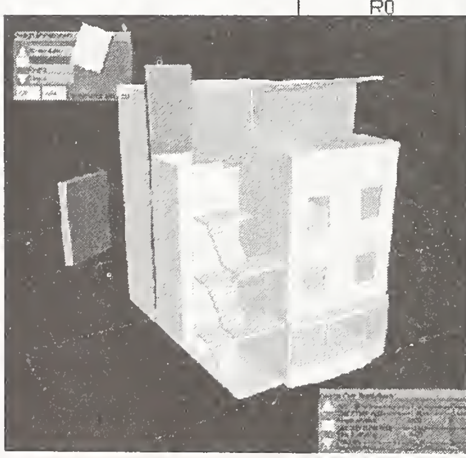


Figure 4. Virtual design environment

Figure 5. Virtual construction site

AUTOMATION IN BUILDING DESIGN WITH SPATIAL INFORMATION

Tang-Hung Nguyen
Assistant Professor, Department of Civil Engineering, North Dakota State University, 120D CME Building, Fargo, ND 58105 – Tel.: (701) 231-8646 – Email address: hung.nguyen@ndsu.nodak.edu.

Amr A. Oloufa
Associate Professor and Director of Construction Engineering Program, Department of Civil and Environmental Engineering, University of Central Florida, P.O. Box 16245, Orlando, FL 32816 – Tel.: (407) 823-3592 - Email address: aoloufa@mail.ucf.edu

Abstract: While current CAD (Computer-Aided Design) systems provide a variety of data representation schemes (e.g. wire-frames, surface, and solid modeling) and data exchange protocols, they fail to address automation issues in retrieval of building information to be used in different AEC (Architecture, Engineering, and Construction) applications. In these CAD systems, topological information (also known as spatial relations) describing the three-dimensional spatial relationships between building components is conventionally represented in a manual fashion into data models. The manual data representation, however, inherently is a complex and challenging task due to the wide variety of spatial relationships. This paper outlines a computer-based building design framework with emphasis on the engine capable of automatically deducing topological information of building components, which support various aspects in building design such as constructability analysis, construction planning, and building code compliance checking.

Keywords: Computer-Aided Design, Topological Information, Automated Building Design Systems, Solid Modeling.

Introduction

Presently, most CAD systems provide two major tools for representing data of building components: the geometric modeling system dealing with spatial abstractions and the database management system dealing with functional information. The spatial information describes the building components' geometry ((i.e. dimensions and locations) and topology (i.e. spatial relations among the components), while the functional data represents all the other discipline-specific properties of the components (e.g. structural, thermal properties, design constraints, and building regulations). Formal representation of spatial data in general and topological data in particular of building components is a complex and challenging task in developing building design systems. The complexity is in part due to the fact that each professional usually utilizes his/her own representation of topologies and dimensionalities to express spatial information of building components. Furthermore, different design tasks require different types of topological information. For example, information about the walls *surrounding* a particular space (e.g. fire zone) is needed for code compliance checking of that space, details of *connections* between individual structural members of a reinforced concrete frame should be provided for reasoning about constructability, and information about the *adjacency* among floors of a high-rise building is required for planning the sequence of construction activities. The variety of topological information to be used throughout the design and construction process may contribute to complexity in data definition, representation and retrieval. Moreover, the topological information of building/product elements is conventionally structured directly in product data models, in which designers are prompted to manually specify spatial information of interest to different disciplines. Such a manual representation is usually subject to data inconsistency, incompleteness, and prone to error, which may result in a misinterpretation of information. Thus, the topological information should be classified and modeled in such a way that the required spatial data for a particular design task (e.g. energy analysis, building code compliance checking, constructability reasoning, and construction planning) can be automatically retrieved. This paper attempts to identify and classifies various topological information commonly used in AEC disciplines into more specific categories such as *adjacency, connection, containment, separation,* and *intersection.* These topological relations are then incorporated into a CAD system so that the required spatial relationships between building components can be automatically deduced to support different design tasks.

Classification of Topological Information

The first step taken in the development of the proposed building design system is to identify various topological information (i.e. spatial relations between building components) commonly used in the AEC domain. The spatial relations are then classified into several main categories to support automatic deduction of building information required by different project participants. Five main categories of the topological information have been identified and named *separation, adjacency, connectivity, intersection,* and *containment.*

- *Separation*: refers to the relation where two building components are physically *separate* from each other. Information about the *separation* relation of a building component with others can be used to check for its code compliance. For example, "Individual roof panels shall be *separated* from each other by a distance of not less than 4 feet measured in a horizontal plane" [BOCA Subsection 2607.2, 1993], implies the requirement of a *separation* between two roof panels.

- *Adjacency*: is introduced to express the type of vicinity that may exist between building components. Depending on the location of one component relative to another one *adjacent-to* it, the vicinity relations may be represented by different terms, such as *next-to* (e.g. mechanical room is *next-to* electrical room), *above, below* (e.g. parking space is *below* lobby area), *right-hand-side,* and *left-hand-side.*

- *Connectivity*: relates to modeling *connected-to, attached-to* or *supported-by* relations between building components. For examples, a suspended acoustical tile ceiling is *connected-to* a concrete frame, a rigid insulation is *attached-to* a protected roof membrane, and a beam is *supported-by* a column.

- *Containment*: corresponds to *contained-in* or *included-in* relations. This type of topological information can describe the spatial relationships between two spaces (e.g. an incinerator is *contained-in* a room space), two building components (e.g. a window is *contained-in* a wall), or between a building component and a space (e.g. an air duct is *contained-in* a ceiling space or plenum).

- *Intersection*: represents *shared-by* relations. A typical application of this type of topological information is that the intersection between a tower crane's boom and an existing building helps verify for constructability of a new building where the available tower crane is selected for material delivery.

The classification of topological relations provides a basis for establishing a building data structure facilitating the organization and management of various topological relations in a building data model. Such a data structure is helpful in making it easy to develop algorithms responsible for the automated deduction of spatial information in the proposed computer-based building design system.

The Automated Building Design System

Basically, the proposed computer-based building design system is built on top of a 3D (three-dimensional) solid modeling system and is capable of automatically deducing various topological relationships between building design objects (e.g. beams, columns, windows, walls, etc.). To achieve such an automated design system, the algorithms for deduction of the five main categories of topological relations described above must be developed beforehand.

Deduction Algorithms

As the proposed building design system is developed in a 3D solid modeler, such solid primitives as vertex, edge, face, cell, and loop are used to describe each building objects. The geometric and topological information about theses primitives can be extracted from the 3D CAD system to deduce the spatial relationships between building objects. Basically, the relationships between vertices and faces representing building components determine spatial relationships between those components. There are three possible relationships between a vertex and a face, i.e. the vertex could lies on the face, to the right or above, or to the left or below the face. This leads to the need for a definition of face outward normal vectors. For this study, all face outward normal vectors are defined by the right hand rule which takes the list of vertices to be counter-clockwise and specifies the outward normal to be the one giving the positive value for points outside the solid objects, as shown in Figure 1.

Figure 1. **Convention for Face Outward Normal**

The building element to be considered for illustrating the deduction of topological information in the proposed system is the space. Depending on the position of a vertex with respect to a face, the value representing the vertex-face relationships can be assigned to –1, 1, and 0, which indicates that the vertex is to the negative side of the face, that is, inside the building space; 1 indicates that the vertex is to the positive side of the face, thus outside the space; and 0 represents those vertices that lie on the face itself.

Example: The algorithm identifies spaces that are *adjacent-to*, i.e. *next-to*, *above*, and *below*, a given space. In other words, the algorithm determines whether the given space shares a common face with any other spaces and all vertices other than those defining the common face are outside the given space. The development of this algorithm is based on several assumptions as follows.

1. All outward normal vectors of faces are defined by the right hand side rule which takes the list of vertices to be counter-clockwise and specifies the outward normal to be the one giving a positive value for points outside the space.
2. All faces bounding a building space are planar and defined by the following equation:
   $$Ax + By + Cz + D = 0$$
3. The values obtained by the substitution of the coordinates of the vertices in the equation of the faces, i.e. Relation Indexes (RI), could be positive, negative, or zero, indicating that the vertex is outside, inside, or on the space respectively.
4. All faces comprising a building space are assumed to be convex. It is noticed that the configuration of a building space can be either concave or convex, which is a critical issue. A simple check can be made to differentiate between the two: if a space has any face that gives a point inside the space with a positive RI, the space is concave; if there is no such a face in the space, it is convex. For simplicity, the building spaces being considered in this work are assumed to be convex.

- Input: S1, S2, and S3 as spaces of a building where S1 ≠ S2 ≠ S3 (see Figure 2)
- Output: S1 is *next-to* S2 and *below* S3
- Algorithm:

1. For every vertex of S2 (or S3) and every face of S1, determine vertices of S2 that lie on a face of S1 (i.e. vertices that give zero RI values with respect to the face). These vertices define a common face between S1 and S2 (or S3).

2. For vertices other than those of the common face of S2 (or S3), compute the RI's indicating their positions with respect to S1.
3. If the RI is positive and the outward normal vector of the common face is parallel to XY plane, then S2 is *next-to* S1. In the case of S3 that also shares a common face with S1 and the outward normal vector of the common face is perpendicular to the XY plane, then S1 is *below* S3 or S3 is *above* S1.

Figure 2. **3D Models of Building Objects**

For more elaborate examples and algorithms to deduce other categories of topological relations, the reader may refer to [Nguyen, 1999].

<u>The Proposed Building Design System: TopoInfo</u>

The algorithms for deducing the five categories of topological information (i.e. *adjacency, containment, intersection, separation*, and *connectivity)* have been implemented into a computer-based building design system, named TopoInfo. TopoInfo is developed on top of AutoCAD Release 14 that provides the underlying geometric solid modeler, in which building components such as columns, beams, slabs, windows, doors, etc. are described as 3D solids. Once the building components have been inputted as 3D solid objects, the basic data including geometry and topology about their primitive representing elements (i.e. vertices, edges, faces, and loops) are created and stored in AutoCAD database. These basic data then can be extracted by means of deduction mechanisms and algorithms to deduce spatial relationships between the building components. This can be done by executing a new command, named CheckTopo, which has been developed using an AutoCAD application development tool (e.g. ObjectARX) and added to the AutoCAD system. The process of the deduction can be summarized as follows:

First, from the AutoCAD drawing editor, the user is prompted to input geometric data and dimension (X-side, Y-side, Z-side, and location) describing the desired 3D building objects. The AutoCAD interface provides high-level and precise mechanisms for manipulating various types of geometric data structures and high-level geometric modeling objects such as prism, cylinder, polygon, line, etc. for creating basic design objects such as columns, beams, slabs, walls ... as shown in Figure 3. These basic objects can be "assembled" to create more complex building objects such as rooms, floors, or construction zones. At the lowest-level data structures, the basic objects are represented as 3D solids by means of primitive elements such as vertices, edges, and faces whose basic information are stored in AutoCAD database. Then, a new AutoCAD command (namely CheckTopo) that is added into the TopoInfo system is called to extract the basic geometric data from the AutoCAD database and check for topological relations between the building objects. This command basically contains a number of functions or methods being developed based on the deduction algorithms described previously. These functions are implemented into the TopoInfo system as a deduction engine, which is responsible for extracting basic 3D geometric information of building objects and deducing spatial relations between these objects. Finally, the output including the deduced topological relations can be stored and retrieved, when needed, to support different tasks of building design such as construction planning, constructability reasoning, and code compliance checking. The output can also be displayed in AutoCAD screen providing designers with a quick feedback on spatial relations (e.g. interference, conflicts) among building components.

Applications of Topological Information

Information about topological relations is essential to different AEC professionals throughout the project life cycle, spanning from building design to construction planning and maintenance. Examples of the applications of topological information to building design and construction, in general, include the following:

- Structural engineers require information about *connections* between individual structural elements for structural performance analysis or constructability evaluation,
- Architects are more concerned about the *adjacency* and *intersection* between building spaces and their boundaries for creation of building shapes and space layouts,
- Building inspectors need information (e.g. thermal properties) of the *surrounding* walls of a particular type of space for building code compliance checking.
- Lighting designers need information (e.g. glazing areas) of windows and/or doors *contained in* exterior walls for day lighting load calculations.
- Construction planners need information about the *vicinity* among construction zones for determining the interdependencies of construction activities to be performed in various spaces, and
- HVAC engineers use topological relations between spaces and their enclosing structures for heat-loss calculations and thermal analysis.

Conclusion

Various types of spatial or topological relationships among building components have been identified and classified into five main categories, i.e. *adjacency, containment, intersection, separation,* and *connectivity* which are essential building information to be repeatedly shared and reused by project participants throughout the design and construction process. This classification is helpful in facilitating the grouping and management of various topological relations in a computerized building data model. The implementation of the proposed building design system, namely TopoInfo, demonstrates the feasibility and practicality of using 3D solid modeling environment to develop a building design system capable of automatically generating complex building information such as spatial relations among building components, which can be retrieved to support different building design activities such as construction planning, constructability reasoning, and code compliance checking.

References

Augenbroe, G.: “COMBINE II: Computer Models for the Building Industry in Europe”, Final Report, Technische Universiteit, Delft, Netherlands, 1995.

Björk, B.: “The Topology of Spaces, Space Boundaries and Enclosing Structures in Building Product Data Models”, the CIB W78 Workshop on Computer-Integrated Construction, Montreal, Canada, 56-92, 1992.

BOCA: “The BOCA (Building Officials and Code Administrators) National Building Code”, 12th Edition, Building Officials and Code Administrators International, Inc., Country Club Hills, IL, U.S.A, 1993.

Fischer, M.: “Automating Constructability Reasoning with a Geometrical and Topological Project Model”, Computing Systems in Engineering, 4(2-3), Great Britain, 179-192, 1993.

Ikeda, M., Oloufa, A.A., and Sekihara, S. “An Object Oriented Environment for the Construction Planning of High Rise Buildings”, Proceedings of the 13th International Symposium on Automation and Robotics in Construction, Tokyo, Japan, 411-420, 1996.

Nguyen, T.H.: “Automated Deduction of Topological Information in a Computer-Based Building Design System”, Ph.D. Dissertation at Department of Architectural Engineering at the Pennsylvania State University, University Park, PA, U.S.A, 1999.

Shaked, O. and Warszawski, A.: "Knowledge-Based System for Construction Planning of High-Rise Buildings", Journal of Construction Engineering and Management, ASCE, 21(2), 172-182, 1995.

ObjectARX: "ObjectARX Software Development Kit for AutoCAD Release 14", Autodesk Inc., U.S.A, 1998.

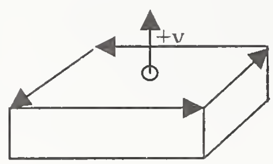


Figure 1. **Convention for Face Outward Normal**

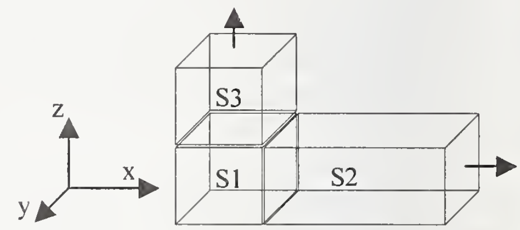


Figure 2. **3D Models of Building Objects**

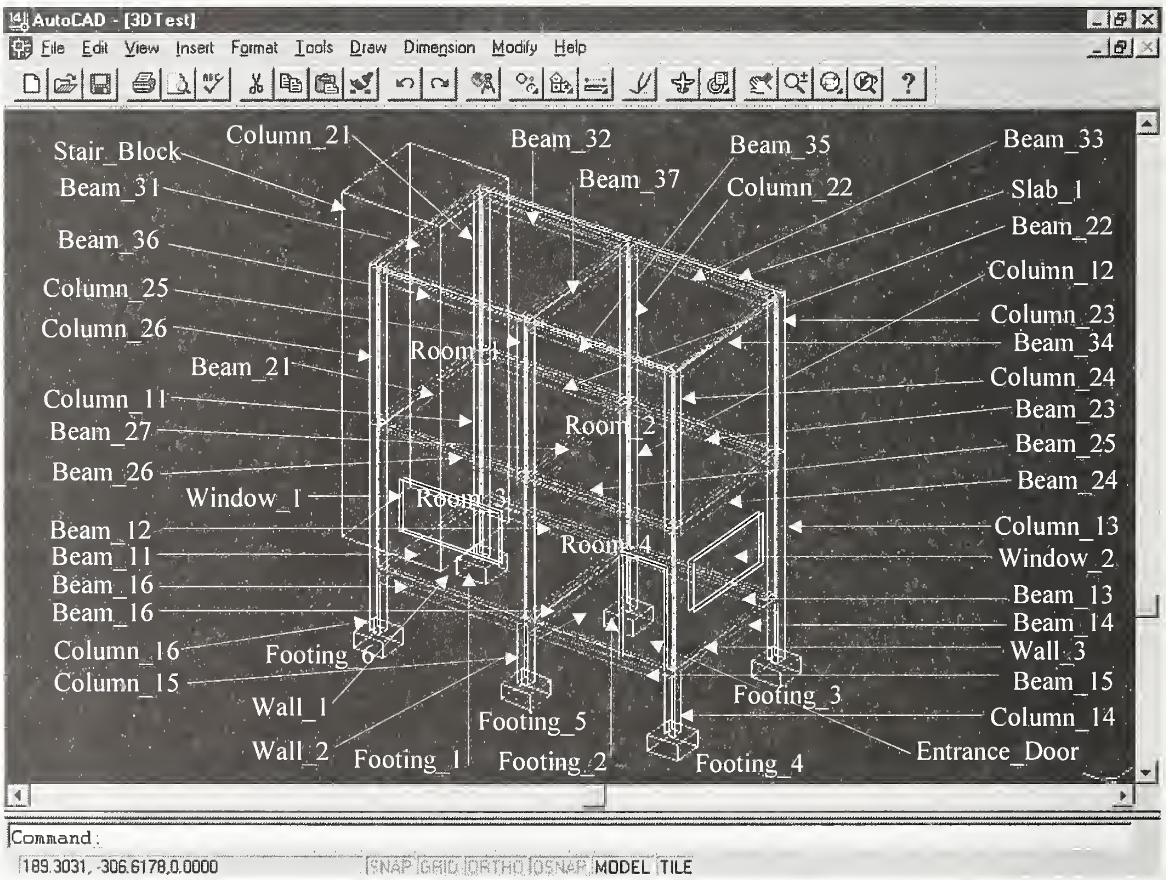


Figure 3. **3D Model of a Typical Building**

GENETIC ALGORITHMS FOR ACCESSING ENGINEERING PERFORMANCE

by

Luh-Maan Chang[1] and Lei Zhang[2]

ABSTRACT

The performance of engineering activities has significant impacts on the successfulness of implementing industrial construction projects. Improving engineering performance can lead to better project outcomes. Previous studies on engineering performance improvement have either focused on the use of certain techniques or products, or looked at specific engineering processes or areas. There has been a lack of a systematic and analytical approach that improves engineering performance based on the understanding of the relationships between engineering inputs and project outcomes. The paper proposes a generic model, which integrates genetic algorithms with artificial neural networks, for modeling engineering performance measurement and improvement in industrial construction projects. Due to their robust and efficient search ability in complex situations, genetic algorithms are employed to search for solutions to improving engineering performance with the searching criteria, fitness function, being the neural networks that establish the relationships between engineering inputs and project outputs.

KEYWORDS: engineering performance, genetic algorithms, artificial neural networks

1. INTRODUCTION

Industrial construction projects have been experiencing unsuccessful implementation of projects for a long time. An industry survey (Post 1998) reported that one-third of the projects surveyed was over budget and nearly half was delivered late. The development of an industrial facility spans over five stages: pre-project planning, detailed design, procurement, construction, and start-up and commissioning (CII 1997). Early researches addressed the impact of engineering performance on the overall outputs of a project. For example, design errors, changes and omissions could constitute approximately 10% of the total installed costs of a project while construction mistakes account for only about 2% (Davis *et al* 1989). 25% of the facility owners surveyed by Post (1998) ranked detailed design as the weak link in the process of facility development.

The Research Team 156 (RT-156) of Construction Industry Institute (CII) studied the industrial project data collected by CII Benchmarking and Metrics Committee. The study reported that the detailed design phase was a prime source of project schedule delays and that about half of the project scope and development changes were initiated during the detailed design phase. The report also pointed out that design errors were the utmost source of field rework and that design-related field rework surpassed that initiated by both owner and constructor (Georgy *et al* 2000).

---

[1] Associate Professor, Division of Construction Engineering and Management, School of Civil Engineering, Purdue University, West Lafayette, IN 47907, 765-494-2246, changlm@ecn.purdue.edu
[2] Doctoral Research Assistant, Division of Construction Engineering and Management, School of Civil Engineering, Purdue University, West Lafayette, IN 47907, 765-494-0696, zhang41@ecn.purdue.edu

Since industrial projects involve huge amount of investment, even a small percentage of cost overrun or schedule delay will result in serious economic loss. Therefore, there is an urgent need to improve project outputs through improving engineering performance. This research aims at searching for approaches to improving engineering performance in industrial construction projects through integrating genetic algorithms with artificial neural networks. Engineering refers to the detailed design phase of an industrial project.

2. PREVIOUS STUDIES ON ENGINEERING PERFORMANCE

Engineering is a systematic process with inputs and outputs. Engineering performance measurement deals with the output side. The ability to successfully perform the engineering and design activities on an industrial construction project depends on various project input variables (i.e., project attributes and conditions), which are essential in driving its engineering performance. There was a lack of analytical scheme that can approximate the cause-effect relationship between engineering inputs and outputs until the research work of Georgy (2000) and CII RT-156, which is part of the foundation of this research study.

CII RT-156 identified a total of 25 engineering input variables and ten engineering performance measures, as shown in Table 1 and Table 2 respectively. A neural-fuzzy system was developed for establishing the relationships between engineering inputs and engineering performance measures. A multi-attribute utility function was used to aggregate the performance measures into a composite index to indicate the engineering performance level (Chang *et al* 2001 and Georgy 2000).

Researchers in the past tried various approaches to improve engineering performance, but most of their approaches are qualitative in nature and have certain limitations. The limitations come from the fact that some approaches promote the use of a specific technique or product and some look at specific areas of engineering and design activities (Armentrout 1986, Atkin and Gill 1986, Breen and Kontny 1987, Choi and Ibbs 1990, Ginn and Barlog 1993). There is a lack of a systematic and analytical approach that looks at improving engineering performance based on the understanding of the relationship between engineering performance and its driving factors.

3. THEORIES

3.1 Artificial Neural Networks (ANNs)

ANNs are an information processing technology that simulates the human brain and nerve system. Their basic element is also called neuron (or node). All neurons are organized in layered structure and connected with weighted links. There is always an input layer where the initial stimulus happens, and an output layer where the final reaction of the system is shot out. ANNs' two major functions are learning and recall. Learning is the process of adapting the connection weights in an ANN to produce the desired outputs in response to inputs. Recall is the process of producing outputs in accordance to specific inputs using the knowledge obtained through learning (Tsoukalas and Uhrig, 1997).

3.2 Genetic Algorithms (GAs)

GAs are robust general-purpose search program based on the mechanism of natural selection and natural genetics (Holland 1972). Genes and chromosomes are the fundamental elements in GAs. A chromosome is a string of genes. In a real problem, genes are the variables that are considered influential in controlling the process being optimized, and a chromosome is a solution to the problem. GAs search for the optimal solution from populations of chromosomes. In this research, the genes are the 25 input variables in Table 1. A chromosome is a set of the 25 input variables. There is an objective function (preferably called fitness function) in GAs. The search process seeks the maximum or minimum value of the fitness function.

4. MODELS

The fundamental approach of the research is to employ GAs to search for the engineering performance inputs that lead to optimal engineering performance. The ANN system shown in Figure 1 serves as a complicated fitness function. Two models were built.

- Engineering Performance Index Model (EPI Model).
- GA-ANN-Integrated Search Model (GA-ANN Model).

4.1 EPI Model

EPI model is in essence the framework of CII RT-156. As illustrated in Figure 1 and Figure 2, EPI model is comprised of two parts. The first part is 10 neural networks that establish the relationships between the 25 engineering inputs and the 10 engineering performance measures respectively. The second part is a multiple attribute utility function that takes the outputs from the 10 neural networks in the first part as its inputs and translates them into a composite utility score, engineering performance index.

The 10 neural networks, after being trained, can predict performance measures for given engineering inputs. The 10 engineering performance measures depict, from different perspectives, the quality of outputs of engineering activities. However, if it is required to evaluate a project or to compare it with another one, it will be hard to make the judgment when 10 varying measures are presented. Therefore, there comes the need for a single composite measure that indicates the overall level of engineering performance and contains the information embedded in the 10 measures. Through multiple attribute utility function, an engineering performance index is defined on the scale of [0, 1] with 0 depicting the poorest engineering performance and 1 the best performance.

Thus, through the trained neural networks, if given engineering inputs, EPI model can make prediction on engineering performance through both a group of 10 different measures and an overall engineering performance index. The set of 10 measures gives a comprehensive view of engineering performance. The engineering performance index will be used as fitness function value in GA-ANN model.

4.2 GA-ANN Model

GA-ANN model, as shown in Figure 3, depicts a typical genetic search process. Its most distinguished feature is the fitness function, EPI model, where the GA-ANN integration happens.

GA-ANN model searches the engineering inputs that lead to better engineering performance. The genetic search starts with an initial population. The initial population is comprised of a number of individuals. Each individual is a chromosome consisting of 25 genes, each of which corresponds to an engineering input in Table 1. For a given project, the input variables related to basic project attributes including general project attributes, general owner attributes and general designer attributes (refer to Table 2) will be kept constant throughout the genetic search; all other input variables subject to the changes in the actual project execution will be manipulated by genetic operations in order to form better combinations of the variables.

GA-ANN evaluates all individuals, keeps the good ones, reproduces the good ones, and sometimes transforms the good ones to make even better ones, ... until satisfactory individuals are produced. First of all, the individuals in the initial generation are evaluated through the fitness function, EPI model. First, Each individual is presented to the 10 trained neural networks that predict its 10 corresponding engineering performance measures. Second, the multiple attribute utility function transforms the 10 predicted measures into a composite engineering performance index, which is the fitness function value of the individual.

Then, the initial generation goes through the genetic operations: selection, reproduction, crossover and mutation. First, the individuals with higher fitness function values get selected and the worse ones eliminated, which means that the engineering inputs that create better engineering performance are kept. Second, the selected ones are reproduced and crossovered. Lastly, a certain percentage of the individuals go through the mutation process which transforms a certain number of genes of the individuals. The mutation process might make the mutated individuals better or worse. Thus, the second generation is formed.

The second generation also goes through fitness evaluation, selection, reproduction, crossover and mutation. Some individuals better than those in the second generation are assembled and come into the third generation. The general trend is that the individuals become better and better from generation to generation. In other words, the level of engineering performance becomes higher and higher.

The genetic search process keeps going on until a certain termination criterion is met. Usually the termination criterion can be a desired fitness value, the maximum number of generations, or computation time. By the time the process stops, one or more sets of engineering inputs will be identified as the ones that lead to an engineering performance level close or equal to the desired level.

4.3 Relationships Between the Models

EPI model establishes the relationships between the engineering inputs and engineering performance measures and aggregates the measures into a composite index to indicate the level of engineering performance. GA-ANN model does the genetic search for better engineering performance using EPI model as the fitness function while EPI model provides engineering performance prediction for given engineering inputs.

5. ANTICIPATED APPLICATIONS OF THE MODELS

For past projects, GA-ANN model and EPI model work together to search better engineering performance and the corresponding engineering inputs. Then, the actual engineering inputs can be compared with the those searched by GA-ANN model and the comparison might be able to indicate what could have been done to achieve better engineering performance.

For future projects, GA-ANN model looks for the possible better engineering inputs and outputs for the project. These anticipated project inputs and outputs might act as the guideline and goal for the actual project execution.

6. DATA ANALYSIS

The project data for validating the proposed models are being collected by the authors. The result of data analysis is expected to be presented at the conference.

7. CONCLUSIONS

This paper proposed a systematic approach to improving the practice of engineering performance. The fundamental idea is to find the possible best practice of engineering activities for a given project. To pursue this, genetic algorithms and artificial neural networks are employed to build the models. Artificial neural networks provide the ability to establish the relationships between engineering activity inputs and engineering performance outputs, and genetic algorithms serve as a search engine to find the possible best engineering practice based on the relationships between engineering inputs and outputs identified through artificial neural networks.

8. REFERENCES

Armentrout D. R. (1986). Engineering Productivity Management and Performance Measurement. *Journal of Management in Engineering* **2:3**, 141-147

Atkin B. L. and Gill E. M. (1986). CAD and Management of Construction Projects. *Journal of Construction Engineering* **112:4**, 557-565

Breen W. C. and Kontny V. L. (1987). Automation's Impact on Engineering Design Progress. *Journal of Management in Engineering* **3:4**, 275-287

Chang L. M.,, Georgy M. E. and Zhang L. (2001). *Engineering Productivity Measurement.* Research Report, RR156-11, Construction Research Institute, Austin, TX

Choi K. C. and Ibbs C. W. (1990). Costs and Benefits of Computerization in Design and Construction. *Journal of Computing in Civil Engineering ASCE* **4:1**, 91-106

CII (1997). Determining the Impact of Process Change on the EPC pprocess: A Model to Quantify Time Cost Improvements. *Report, IR125-2*, Construction Industry Institute, Austin, TX

Davis K., Ledbetter W. B. and Burati J. L. (1989). Measuring Design and Construction Quality Costs. *Journal of Construction Engineering and Management ASCE* **115:3**, 385-400

Georgy M., Chang L. M. and Walsh K. (2000). Engineering Performance in Industrial Construction. *Proceedings for Construction Congress VI*, ASCE, Orlando, FL

Georgy M. (2000). "Utility-based Neurofuzzy Approach for Engineering Performance Assessment in Industrial Construction Projects". Ph.D. Dissertation, Purdue University

Ginn D. and Barlog R. (1993). Reducing Complexity and Working with Bottlenecks Improves an Oil Refinery's Engineering Performance. *National Productivity Review*, Winter 1993/1994, 63-77

Holland J. (1975). *Adaptation in Natural and Artificial Systems*, The University of Michigan Press

Post N.M. (1998). Building Teams Get High Marks. *Engineering News Record* **240:19**, 32-39

Tsoukalas L. H. and Uhrig R. E. (1997). *Fuzzy and Neural Approaches in Engineering*, John Wiley & Sons, Inc.

Table 1 Engineering Input Variables

Category	Variables
General project attributes	Project size (total installation cost)
	Contract type
	Relative size of project compared to projects of the same industry type
	Relative level of complexity
	Site conditions
	Legal and environmental conditions
General owner attributes	Owner profile and participation
	Newness of process technology to owner
	Owner previous experience with designer
General designer attributes	Split engineering practices
	Designer qualifications and capacity
	Newness of process technology to designer
Project schedule	Design schedule
	Design-construction overlap
Project information inputs	Completeness of scope definition
	Completeness of objectives and priorities
	Completeness of basic design data
	Quality of constructor input and constructability
	Quality of vendor data
Level of automation	Use of 3D CAD modeling
	Use of Integrated Databases (IDB)
	Use of Electronic Data Interchange (EDI)
Project changes	Percent TIC scope changes
	Change management procedure
	Change communication system

Table 2 Engineering Output Variables (Engineering Measures)

Category	Variables
Detailed design value	% design rework
	Design document release commitment
	% detailed design schedule delay
	% detailed design cost overrun
Fabrication and construction value	% fabrication and construction schedule delay due to design deficiencies
	% fabrication and construction cost overrun due to design deficiencies
	% construction hours for design problem solving and field design
	% estimated dollar savings due to constructability
Start-up and commissioning value	% start-up schedule delay due to design deficiencies
	% start-up cost overrun due to design deficiencies

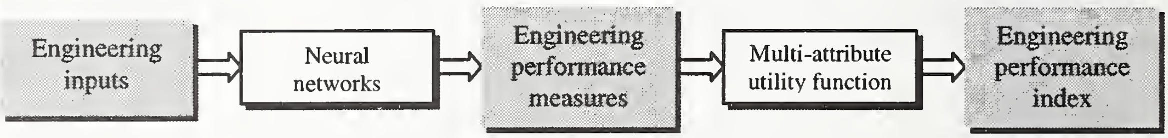


Figure 1. EPI Model -- General Idea

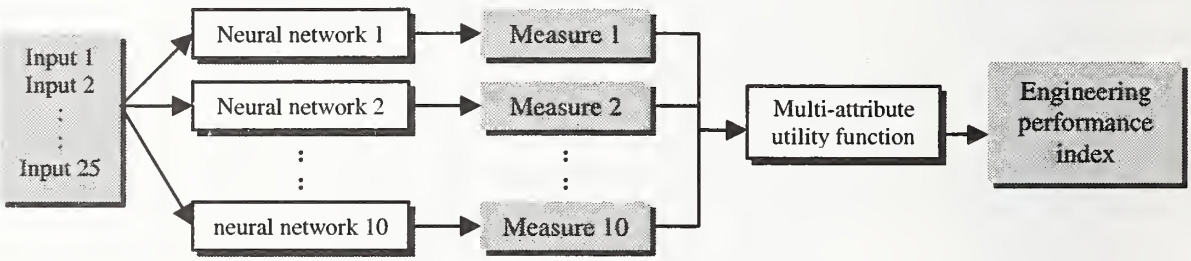


Figure 2. EPI Model -- Breakdown

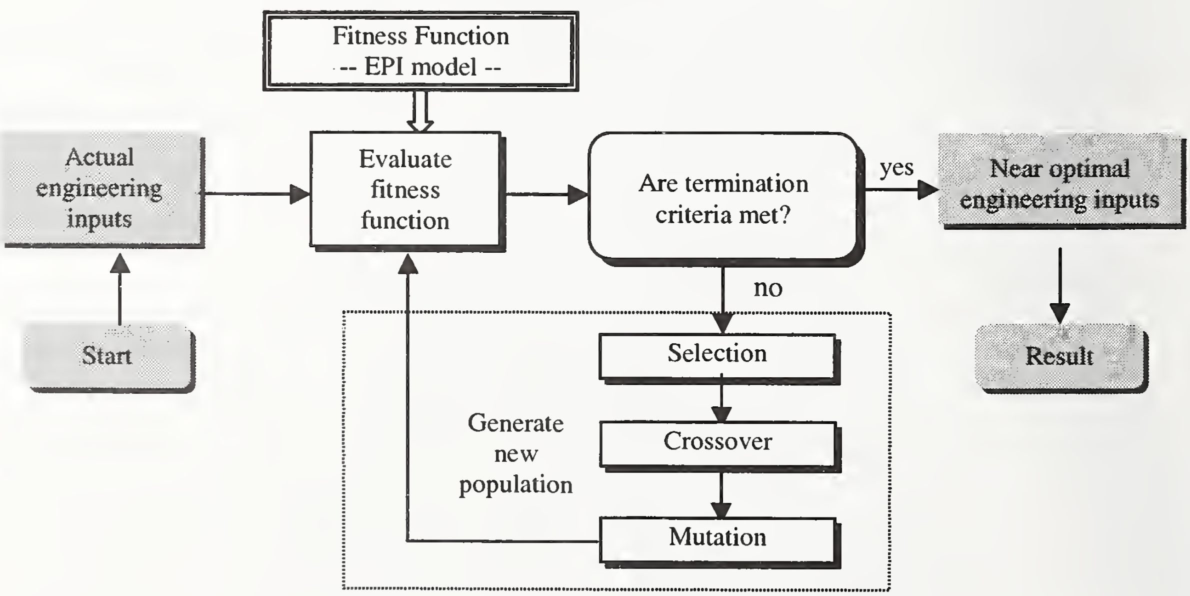


Figure 3. GA-ANN Model

Mixed Reality Benefits For Design Perception

by

Phillip S. Dunston[1], Xiangyu Wang[2], Mark Billinghurst[3], Ben Hampson[4]

[1] *Construction Engineering and Management Division*
School of Civil Engineering
Purdue University
West Lafayette, Indiana
(dunston@ecn.purdue.edu)
[2] *Department of Civil and Environmental Engineering*
University of Washington
Seattle, Washington
[3] *Human Interface Technology Laboratory*
University of Washington
Seattle, Washington
[4] *The McKinstry Company, Seattle, Washington*

ABSTRACT: Design visualization is key to the communication and shared perception of designs and is essential for meaningful design development and collaborations. The initial development of an Augmented Reality Computer Aided Drawing (AR CAD) system for enhancing visualization of models created in standard CAD was presented at the 17th ISARC. AR CAD features a more natural mode for changing views of the model and completely understanding the model content. Expected benefits are improved efficiency in the design detailing function, both for the individual detailer and for design collaborations where maintaining an accurate shared understanding of the design model is critical. An experimental program is under way to examine the impact of AR CAD upon a user's perception and recall of a design model. Related experiments with desktop and immersive virtual environments have found that motion cues can indeed markedly improve spatial cognition. It is expected that we will see the same benefits in our AR CAD system, although until now such studies have not been conducted in an AR environment. This paper presents the rationale for experiments to measure the impact of AR CAD in terms of cognition cost, and it lays the foundation for further application of Mixed Reality (MR) technology to the design, construction, and maintenance phases of a facility's life cycle. MR applications may prove promising for effective communication of designs for prefabrication, site installation, and the planning and excecution of maintenance operations.

KEYWORDS: 3D CAD; Augmented Reality; Mixed Reality; spatial cognition; visualization; Virtual Reality

1. INTRODUCTION

Recognizing 3D CAD and walk-through software as the state-of-the-art for visualizing design details, Augmented Reality Computer Aided Drawing (AR CAD) was introduced at the 17th ISARC [1]. In addition to providing an alternative three-dimensional view of construction models, it was proposed that experimentation with the spectrum of Mixed Reality as illustrated by Milgram and Kishino [2] in Figure 1 could open the door to more modes of interaction with design content than is currently available through typical CAD software. Alternative interface metaphors can be developed and tailored to facilitate development and communication of designs. Thus, our work is aimed at determining the appropriate MR modes for planning, design, construction, maintenance and the associated interface tools.

Augmented Reality (AR) occupies that place in the continuum where virtual objects are inserted to a predominantly real world scene. AR also allows virtual enhancements to physical interface objects. MR offers a broader range of options for interfacing with digitally based information. In our present work we are interested in applying more innovative AR techniques to the design phase of the construction process, the AR CAD system is designed for exploring the benefits of supporting design, and ultimately construction, with various modes of Virtual Reality interfaces.

The AR CAD system utilizes fiduciary markers in the real world to position the model and allow the user to easily see it from any viewpoint. As such, it offers the benefit of a very natural mode for changing views of the model and completely understanding the content than would be afforded by visualization systems that have a more constrained means of navigation. This feature is expected to improve efficiency in the design detailing function, both for the individual detailer and for design collaborations where maintaining an accurate shared understanding of the design model is critical.

An experimental program is underway to examine the impact of AR CAD upon a user's perception and recall of a design model. Related experiments with desktop and immersive virtual environments have found that motion cues can indeed markedly improve spatial cognition. It is expected that we will see the same benefits in AR CAD, although until now such studies have not been conducted in an AR environment. Positive results from these experiments would confirm that AR CAD has the potential to support the reduction of errors during design detailing and the more rapid and effective resolution of space conflicts interferences during design collaborations. This stage of research lays the groundwork for further application of MR technology to the design, construction, and maintenance phases of a facility's life cycle. MR applications may prove promising for effective communication of designs for prefabrication, site installation, and the planning and execution of maintenance operations.

This paper presents the updated description of the features of the AR CAD system and also explores some spatial cogniton issues that may arise associated with the system. We also explain the rationale for experiments to determine the benefits of AR CAD over standard CAD (AutoCAD in this instance) in conflict detection. Such experiments constitute our first attempt to evaluate AR CAD with regard to spatial cognition issues.

2. AR CAD SYSTEM

The AR CAD system has been modified for improved performance over its predecessor version. The current experimental facility can provide the piping detailer with the ability to explore the CAD design in the non-immersive (AR) virtual reality mode and still consists of the following components as first described by Dunston et al. [1] and shown in Figure 2.

Modeling Computer: The modeling computer runs AutoCAD and a specially designed AutoCAD plug-in. The CAD detailer designs the model on this machine and then sends the 3D model information out to the AR computer using the AutoCAD plug-in. The plug-in software uses standard network communication code to communicate between computers. Communication can also be between modules in the same computer. The current system functions on a single desktop Pentium 4 PC with 1.6 GHz.

AR Computer: The AR computer runs an Augmented Reality application that allows a user to see virtual 3D models superimposed over a real-time video-recorded view of the real world. The AR application is a custom application that is based on the ARToolKit library [3] and OpenGL library [4] and a database, containing simplified 3D models of pipe structures in the VRML file format. The AR application receives the 3D model information through the network communication and then instantly creates a 3D virtual model of the design.

Camera: Also connected to the AR computer is a Logitech QuickCam Pro 3000 Camera with video capture of up to 640*480 pixels and frame rate of up to 30 frames per second. The computer performs image processing on the video image from the camera to find specially marked tracking cards. The camera's position can be calculated from a tracking card and a virtual model overlaid on the card. The resultant composite image is fed back into the desktop display for the user to see. The result is a view of the real world scene with a 3D virtual model overlaid on it (Figure 3). This tracking technique enables the user to easily view the model from any perspective above the card.

3. NEW FEATURES IN THE SYSTEM

Several feature modifications or additions have been made to the AR CAD prototype. These features are as follows:

1. New Graphics Library: The previous version of AR CAD relied upon the LibVRML97 library for the rendering of virtual images while the newer version relies upon the OpenGL standard for greater versatility and generation of more stable virtual models. OpenGL is a graphics library that is less memory intensive and facilitates real-time interactions with the virtual models.

2. Automatic Conflict Detection: The program will automatically detect any conflict or interference appearing among the pipe objects. If there is a conflict, the interfered objects will be highlighted as wire frame elements on the screen rather than the default solid model representation (see Figure 4). This task is accomplished using the boundary box feature of OpenGL. A similar conflict detection function has been developed by Shiau et al.[5] whose application uses ellipses to identify interferences in structures, appliances and piping systems in 3D models.

3. Objects Selection and Manipulation: If a certain object is selected (activated), the wire frame of it will appear on the scene, which makes the designer easily recognize which object is activated so that the user can use the mouse to move and scale any object activated and also use the keyboard to rotate the object along local x, y, z axes. Brief information describing the selected object will be shown as a text string in the bottom of the screen.

A zooming feature has also been added to the AR module (see Figure. 5). Another potentially beneficial function under development is a transparency mode. This rendering mode can provide distant objects with a degree of visibility even if the view is obstructed by nearer objects. Finally, we are in the midst of adding a fly-in feature that will provide an immersive VR view of the design space.

4. SPATIAL COGNITION COST

Since AR CAD presently acts mostly as an assistant viewing tool for standard AutoCAD, the chief issues that have arisen are those concerning human spatial cognition. While 3D modeling is readily accepted as being less abstract and therefore an intrisically more meaningful graphic form for representing and communicating complex designs, there are still open questions due to the degree of separation from reality that yet remains, questions relating to control of perspective and user interface metaphors. Some features or options of the AR CAD software raise the question of costs and benefits of AR CAD with respect to spatial cognition. That is, do the spatial cognition benefits outweigh the cognition cost associated with a viewing assistant mode?

Very limited research has been done on spatial cognition issues associated with AR applications. However, some testing results done by other researchers indicate that subjects are able to acquire configuration knowledge of immersive virtual environments in spite of the fact that the subjects lack the benefit of spatial calibration derived from physical movements through a real environment. In this paper, we are concerned with the time cost of necessary cognitive processes.

During construction of spatial mental models, switching perspective from one scene to another scene exacts a cost [6]. Cognition cost is a kind of cognitive cost incurred by mental transformations associated with the such changes as a new referent object or frame, or change of a viewpoint while switching perspectives. Perspective switching due to transitions between AutoCAD and the AR scene will incur a cognition cost (see Figure 6 for comparison of AR CAD model and AutoCAD model). However, presently, we don't know how much spatial cognition benefit—a more natural and more smooth navigation between viewpoints—AR CAD can provide versus the cognition cost of using the system as we have designed it to presently operate.

To understand what that cost may be, we need to analyze what perspective switching entails. A perspective consists of a referent object or frame and a viewpoint. Each of these components changes when perspective is changed. Each of these changes requires different mental transformations and cognition cost associated with them. However, the relative costs of the transformation are not yet fully known.

The method for changing viewpoint in the AR scene is quite different from that used in AutoCAD. For AR scene navigation, the viewpoint is embedded in real world background and keeps changing as the user changes the relative position between the camera and the tracking marker. In contrast, standard CAD software like AutoCAD has a single color background with only a simple symbol referencing the coordinate system origin. There is nothing else in the scene to which the user can reference his or her viewing perspective. Furthermore, the mechanisms for changing views is often not intuitive.

Even though the relative position of each object in the AR scene is absolutely the same as the one in AutoCAD, another significant cognition cost will come from additional mental processing demands for reconciling misaligned headings. When transitioning between scenes (AR and AutoCAD) in which models are misaligned, mental calculations are required to account for the different headings in each scene. This condition is very similar to alignment effects found for map usage [7]. If an additional step is necessary to compute the direction of a location for misaligned headings, there is a resulting processing demand and cognition cost.

While there is indeed a cognition cost associated with switching scenes or perspectives, in some cases, the costs of switching may not be greater than the cost of staying with the same scene. One of the major advantages of AR CAD is a natural and smooth transition between viewpoints making it easier to locate any point or corner of the virtual scene quickly. However, the big drawback internal to AR CAD is the cognition cost in transitioning between the virtual scene in AR and the original model in AutoCAD. There is a trade-off in utilizing these two viewing environments together. However, right now, we do not know how much spatial cognition benefit is associated with AR CAD, nor how much cognition cost. If a specific benefit is identified, then the cognition cost for obtaining it can be measured for standard CAD (AutoCAD) alone and for CAD equipped with the AR viewing mode. This approach can be used to validate any sort of viewing assist mode or function. For instance, conflicts or interferences can be identified in the AutoCAD view by means of certain visible features or by transitioning to the AR scene to use the automatic conflict detection feature and then returning to AutoCAD to make corrections. The simple measurement of time indicates the relative cognition cost. We are performing such tests, incorporating statistical design to allow for differences between user subjects and model complexity. Preliminary trials have indicated the AR viewer assist to be a worthwhile means for identifying conflicts without increasing a detailer's overall time in detecting conflicts.

5. CONCLUSION

This paper has presented an update on the development of an Augmented Reality viewer assist feature that turns standard CAD into AR CAD and has discussed the importance of considering spatial cognition issues in the validation of this system. Future work will pursue confirmation of AR CAD benefits for design model perception for both individual users and collaborating partners. These efforts will also include feature enhancements to extend AR CAD's applicability, such as more seamless generation of the virtual models. The long term objective is the validation of Mixed Reality as a useful technology arena for effecting human interfaces with digitally based project design information.

6. ACKNOWLEDGEMENTS

The authors are grateful to the McKinstry Company of Seattle, Washington (U.S.A.) for their continuing cooperation in this research.

7. REFERENCES

1. Dunston, P.S., Billinghurst, M., Luo, Y. and Hampson, B. 2000. "Virtual Visualization For The Mechanical Trade." *International Symposium on Automation and Robotics in Construction (ISARC),* National Taiwan University, pp. 1131-1136.

2. Milgram, P. and Kishino, F. 1994. "A Taxonomy of Mixed Reality Visual Displays." *IECE Trans. on Information and Systems* (Special Issue on Networked Reality), Vol. E77-D, No. 12, pp. 1321-1329.

3. ARToolKit information website. http://www.hitl.washington.edu/research/shared_space/

4. Woo, M., Neider, J., Davis, T, and Shreiner, D. 1999. OpenGL Programming Guide, *3rd* edition, OpenGL Architecture Review Board, Addison-Wesley Longman, Inc.

5. Shiau, Yan-Chyuan, Wang, Ming-Teh and Yang, Chung-Ham. 2001. "The Application of 3D and Interference Detection in Construction Engineering." Proceedings of the *18th International Symposium on Automation and Robotics in Construction (ISARC).* E.Bundy and J. Szlagowski, editors, pp. 95-100.

6. Tversky, Barbara, Lee, Paul and Mainwaring, Scott, 1999. "Why do speakers mix perspectives?" *Spatial Cognition and Computation* 1(4), pp. 399–412.

7. Levine, M., Jankovic, I.N. and Palij, M., 1982. "Principles of Spatial Problem Solving." *Journal of Experimental Psychology General* 111(2), pp. 157–175.

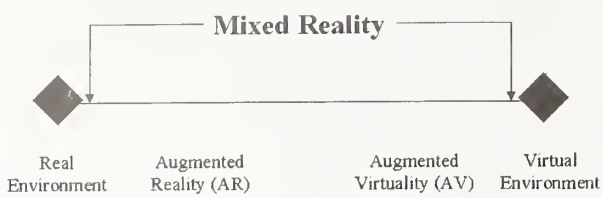


Figure 1. Mixed Reality encompasses all combinations of virtual and real.

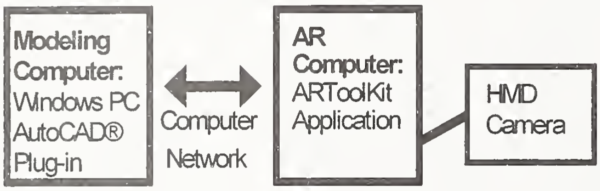


Figure 2. Components of the AR CAD prototype can operate on separate or a single computer.

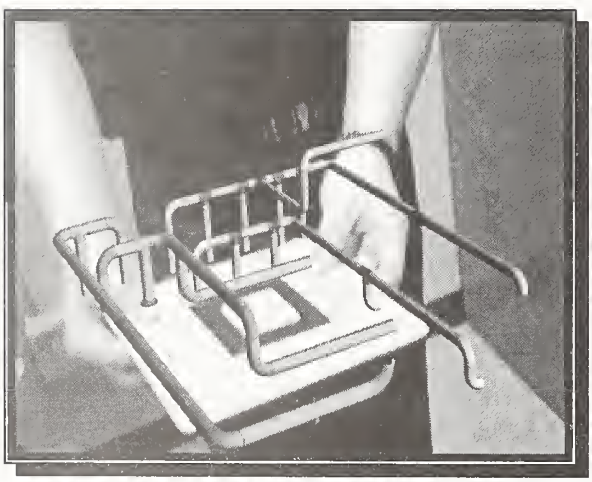


Figure 3. The real world scene is overlayed with a 3D virtual model.

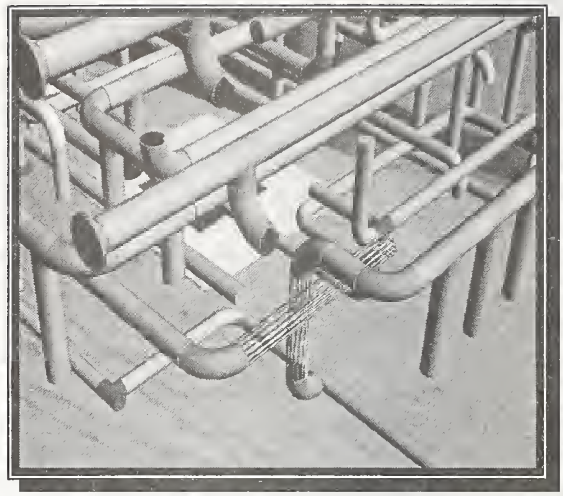


Figure 4. Wire frame representation identifies object interferences.

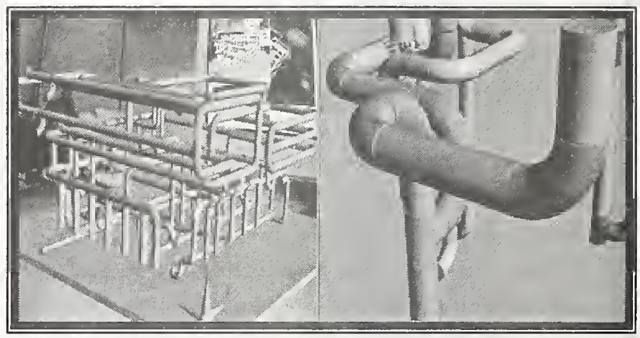


Figure 5. A zooming feature allows close up inspection.

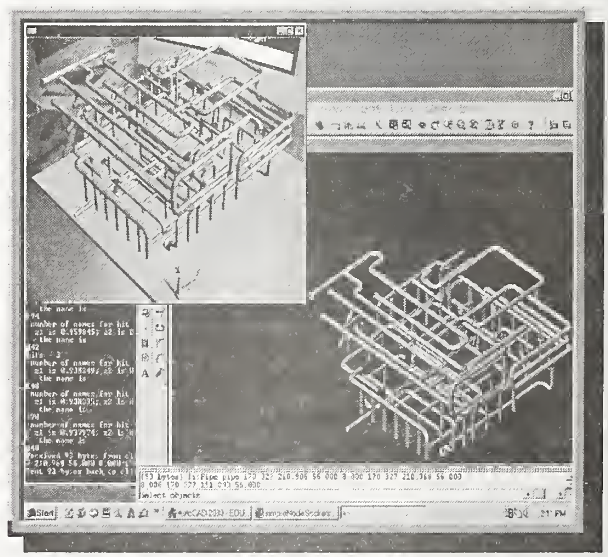


Figure 6. Views in AR and AutoCAD are not usually so aligned when switching scenes.

AUTOMATION CONSIDERATION DURING PROJECT DESIGN

by

MATT M. HEWITT
JOHN A. GAMBATESE

ABSTRACT:

Automation as pertaining to the construction industry is the use of mechanical and electronic means to achieve automatic operation or control to reduce potential exposure, time, or effort while maintaining or improving quality. Contractors utilize automated technologies on projects as a means of saving cost, reducing project durations, improving quality and consistency, and gaining other related project benefits. Communication between the constructor and designer of the construction means and methods to be used is often limited as a result of contractual relationships and competitive bidding requirements. This commonly leads the designer to assume conventional construction equipment will be implemented rather than specialized automated technologies. For this and other reasons, designer consideration in a project's design of the use of automated construction technologies is limited. This paper describes a study to investigate the ability of designers to consider the use of automated construction technologies in the design of a project. The study identifies design practices that facilitate the implementation of automated technologies and exposes barriers, within both the design process and the overall project development process, to the consideration of automation in the design. The findings of the study can be used when one is considering the implementation of construction automation technologies during the design process.

KEYWORDS: automation, design, construction, constructability, equipment, robotics

1. INTRODUCTION:

Over the past few decades, improvements to productivity in the construction industry have been insignificant compared with other industries. Productivity improvements in other industries, especially manufacturing, have stemmed to a large extent from the effective implementation of new technologies. The introduction of new technologies in the construction industry to fully automate the building process has been limited. The same is true for heavy/highway construction. The construction industry remains a craft-oriented and labor-intensive industry with minimal automation of tasks.

The lack of automation in the construction industry can be attributed to many factors. One of the hurdles to automating the construction process is the design of a project. That is, the design of a facility inhibits both the use of available automated equipment during construction and the successful development of new automated equipment. Furthermore, the capabilities of automated equipment are constrained by the physical aspects of the design. Minor modifications to designs can potentially enhance the use of automation and lead to increased construction productivity.

This topic was the focus of a research study to investigate the design practices that facilitate construction automation. This paper describes the study efforts along with identified design practices that enable construction automation and barriers to designing for automation that were exposed.

Department of Civil, Construction and Environmental Engineering
108 Apperson Hall
Corvallis, Oregon 97331-2302
Phone: 541-737-8913
E-mail: john.gambates@orst.edu

2. RESEARCH METHODOLOGY:

The goal of the research study was to improve the ability to prepare designs that facilitate the use of automated technologies in construction work. Designing for construction addresses the constructability of a project and, in this case, a subset of a project's overall constructability - "automation constructability". To meet this goal, the research effort included two objectives. The first objective was to develop and accumulate recommended design practices which, when implemented, enhance the ability to automate construction activities. The recommended practices reflect the capabilities and limitations of current automated technologies and the application of current design practices. Secondly, barriers to the use of automated technologies in the construction process as a result of design features, methods, and deliverables were identified. Knowledge of the barriers provides additional guidance in both the planning and design processes.

Literature accessible through libraries and the World Wide Web was reviewed to determine relevant automated technologies available to the construction industry. This includes technologies which were under research and development, but when ready would be available to the construction industry. Automated technologies were found world wide and filtered based on relative potential for use in the United States construction industry. Technologies were categorized according to the finished product which they are used to create.

To determine industry usage of automated technologies, a survey was sent to the top five hundred contractors as ranked by Engineering News Record (ENR). Companies were asked to respond to several specific questions concerning industry perception, benefits, barriers, etc. of implementing automated technologies. The survey included several questions specifically written to determine whether constructors found value in implementing automated technologies.

Additional information was obtained through interviews of industry personnel. Contractors, designers, fabricators and personnel in other industries were interviewed. Standardized interviews were conducted through structured questions concerning the benefits, barriers, and limitations of automated technologies.

The contractor interviews were conducted from the perspective of both the project manager and equipment operator. Careful considerations were made to obtain a range of companies, based on annual revenue. Design firms were selected based on the type of designs produced and employee count was used to determine their relative size. Fabricators were also interviewed to get another perspective concerning automation. Interviews with fabricators were based on the grouping of technologies available to constructors and the potential for influence by the fabricator. Recognizing that automation exists in many other industries, interviews of personnel in other industries were also conducted to include information concerning how well other industries have come to include automated technologies.

3. RESULTS:

The survey of top contractors provided a list of available automated technologies and questioned whether each firm uses the technologies, has considered using the technologies, would recommend using the technologies to others, and if there is a perceived value in developing the technology. The survey intent was to determine or attempt to quantify the level of automated technology use by the industry. The survey revealed that there is a strong prevalence in concrete and masonry construction of automated mobile screeding, rebar bending, and concrete surface treatment. In addition, responses suggest that there is value in developing further technologies for automating reinforcing cage fabrication for beams and columns, and for mobile bricklaying. Excavation and demolition responses indicate that remote control soil compaction is used exceedingly more than other technologies within the category. Automated dump trucks and global positioning system (GPS) aided excavation appear to be highly considered applications for implementation, but with little or no actual implementation. One such implementation has proven to reduce cost by thirty percent through the use of automated trenching and pipe placement (Lee, Lorene, and Bernold 1988). Inspection and surveying activities show strong prevalence of robotic pipe inspection; otherwise technologies have not yet been greatly implemented for these tasks. For structural steel construction, metal welding and cutting systems are widely used and are recommended to other industry members. Also, robotic welders are implemented extensively amongst industry members. Lastly, material and asset management technologies appear to have developmental value amongst industry members.

Interview surveys consisted of questions concerning best practices, barriers, and changes to the design process or project design. The following summarizes the survey results.

3.1 Contractor: PM & Equipment Operators

According to project managers, over seventy percent of contractors interviewed use some form of automated technology in construction and over seventy five percent have considered the use of automated technologies. Also, implementation is attributed predominately to cost, followed by production, and then quality. Equipment operators perceive production, cost, and quality as benefits of automation and have found through experience that production, cost, schedule, and quality are benefits of implementation. Project managers are unified in agreement that having constructors and designers working together to define automated technologies that needed to be developed would be of value.

3.2 Designer:

The survey results indicate that designers typically do not formally consider construction automation, but when it is considered, the decision is based predominately on cost and quality. Responses also suggest that not many designers are aware of the automated technologies available to contractors, nor does there seem to exist any reference material for individuals seeking information concerning automated technologies and design. While no resources may commonly be used at this time, it should be noted that there exists a website where designers can go to gain information concerning emerging construction technologies (www.new-technologies.org/ECT/Index.html).

3.3 Fabricator:

According to the fabricators interviewed, design for automation is based on cost and quality, followed by productivity and safety. Similar issues to those mentioned previously arise in the fabrication of products and again there is a great deal of value placed on the communication between the designer and the manufacturer/fabricator.

3.4 Best Practices

Accumulated from the literature review and surveys were recommended design practices to facilitate the use of automation during construction. The practices can be categorized as project-related and industry-wide. Project-related recommended practices are as follows:

- Conduct constructability reviews that incorporate consideration of construction automation.
- Standardize building elements.
- "In order to optimize the use of automated technology, it is important that design principles based on the technology are considered (Howe, 1998)."
- Provide adequate clearance for automated technologies to operate.
- Prioritize design objectives (cost, quality, safety, etc.) and compare design alternatives.
- Use electronic documents and make them available to the contractor.
- Consider the capabilities and limitations of the automated technologies. Table 1 provides examples of design practices related to various construction activities in which automated technologies and equipment are notably evident.

With respect to the construction industry as a whole, it was recommended that the technologies be marketed to designers and owners. Primarily the technology manufacturers would undertake this with assistance from constructors. Many designers and owners simply do not have enough exposure to projects that implement automated technologies and, due to this aspect, there is no benefit acknowledgement concerning implementation of automated technologies. Such marketing efforts would expose the designers and owners to the capabilities and benefits of construction automation.

3.5 Barriers

The study revealed numerous barriers to the consideration of construction automation in the design phase. As with the recommended design practices, the barriers can be separated into those that exist at the project level and those are present industry-wide. Project-level barriers are as follows:

- Automated technology capabilities are limited and create tremendous costs when contractors attempt to match project variability with automated equipment variability.
- Frequent changes/advances in the technologies. Technological advancement and improved design occur rapidly and many users cannot keep up with the changes.
- The cost of owning and operating automated technologies.
- A lack of standard design elements. Repetitious elements are likely to lead to greater utilization of automated technologies.
- "No two sites present the same problems, and a layout suitable for one site maybe quite unsuitable for another (Cusack 1994)."
- A lack of consideration of the construction phase by the designer, due to the means and methods residing with the constructor.

Other barriers exist that are not necessarily applicable to a specific project, but are representative of the nature and structure of the construction industry. These industry-wide barriers are as follows:

- Designers typically have limited construction experience. There should be pre-construction consulting between the designer and constructor concerning cost saving construction methods.
- Designers typically have limited knowledge of automated technologies. There should exist consulting between designers and constructors concerning automated technology availability and potential implementation.
- Designers typically have limited knowledge of the design practices, which facilitate the use of the automated construction technologies.
- A general lack of designer interest in considering automated technologies in their designs.
- The structure of, and the roles associated with, the traditional design-bid-build contracting method. Barriers exist which limit the amount of pre-construction consulting and communication that can occur.
- The traditional roles and responsibilities of the designer. Traditionally the means and methods of construction have been the responsibility of the constructor. Consequently, a designer's ability to influence the implementation of construction methods has gone unused or unrecognized.
- Resistance to change from the commonly used design practices.
- Some designers view change as high risk, because there exists a level of uncertainty and untested consequences. Particularly risk associated with implementing automation.
- Conflict of interest on public projects when contractors are brought in during design. Publicly funded projects are required to be competitively bid, limiting the pre-construction communication between constructors and designers.

- Lack of reference material available for designers to use for consulting.

4. EXAMPLE OF AUTOMATION BENEFITS

The benefits of construction automation, and the design impacts on the use of construction automation, can be illustrated using a past project on which a concrete extruder was utilized. The project, located near Clackamas, Oregon, involved the construction of a freeway overpass. Nearly three thousand feet of MSE retaining wall rails were constructed during the two-year project duration, but not all of the railing was extruded. The only railing that was extruded was railing attached to wall "H". Wall H is over 2000 feet long and the extruder contractor completed the wall in two days. Although wall H contained the longest stretch of railing, there were other opportunities to utilize an extruder. Railing extending along another wall, wall "P", and onto the bridge itself, was less than 1000 feet in length and took the bridge contractor over three weeks to complete. Working crews for both walls were relative in size. The railing along wall P and the bridge was not extruded, but could have been. The extruder contractor pointed out that the railing on wall H had a vertical back and curved face, allowing for optimal utilization of their extruder, while the wall P and bridge railing did not. Also, there was a height difference between the railing on wall P and the bridge railing. Although attached, the height transition added to the overall cost of extruding. Since the wall P and bridge railing was under three hundred feet and required a change in shape configuration, the extruder contractor could not competitively compete against conventional construction methods. Also, the railing on wall P and the bridge had two vertical faces, which magnifies any inconsistency and adds to the cost of using an extruder. The extruder contractor added that having at least one curved face helps to reduce visibility of inconsistencies to the human eye.

5. CONCLUSIONS:

Implementation of automated technologies can be greatly affected through the design process when consideration is given to the contractor's opportunities for automation applications. This is made feasible through fluid communication between the constructor and designer. This communication should optimally take place during the design phase. If the designer is not already knowledgeable about the construction means and methods to be used, it is beneficial if the constructor conveys the means and methods for constructing the project to the designer and at a minimum provides some evidence of the automated technologies being considered for implementation. In addition, owners need to perceive the construction process in its entirety, starting from the architect/designer and completing in the hands of the constructor. By holding each entity accountable for cost, quality, and safety, owners can indirectly influence the driver for designers to incorporate designs that will lead to potential cost saving automation and increased project safety.

6. RECOMMENDATIONS:

Designers and constructors are encouraged to take a non-adversarial approach to communication and information exchange with digital documentation and plans, as well as intentions for methods of construction and project follow through, on the behalf of the designer. Lastly, constructors and designers are encouraged to develop business relationships with one another that foster competitive, low cost, high quality, safe projects and in doing this automation implementation can be better achieved.

REFERENCES:

1. A. Scott, Howe, "Designing For Automated Construction", 14th International Symposium on Robotics in Construction, 1998.
2. Emerging Construction Technologies, www.technologies.org/ECT/Index.html.
3. Lee, Junbok; Lorene J. Steven; and Bernold E. Leonhard. "Saving Lives And Money With Robotic Trenching And Pipe Installation", *Journal of Aerospace Engineering.* New York, N.Y. 1988.
4. M. Cusack, "Automation and Robotics." The Industrial Robot, International Fluidics Services, ltd. 1994.

Table 1. Example Design Practices to Facilitate Construction Automation

Automation Category	Example Design Practices
Concrete Finishing	▪ Tighter specifications for grade, slope, and smoothness. ▪ Long continuous paving with constant slope. ▪ Keep floor utilities together, not spread out. ▪ Pour constrained to machine width; no perpendicular curvature. ▪ Eliminate mid-slab protrusions.
Concrete Reinforcement Fabrication and Placement	▪ Use repetitive sizes throughout building for columns, beams (i.e. Revise reinforcing and/or concrete strength in lieu of changing member size.). ▪ Use round columns vs. square columns. ▪ Conducive standards for local iron manufacture facility. ▪ Control degree of architectural variability.
Earthmoving / Excavation	▪ Digital copy of existing surface profile. ▪ Uniformly sloped grading plans with well-defined break-lines. ▪ Precise standards for locating underground utilities effectively. ▪ Depths conducive to the use of automated equipment. (Stronger pipe doesn't get buried as deep, so automation implementation may balance or lower costs through increased productivity)
Soil Compaction	▪ Allow for adequate width of equipment. ▪ Use backfill material conducive to the equipment. ▪ Adjust trench slopes to slope range of equipment. ▪ Adjust lift specifications to allow clearance for perpendicular objects. (Pipes, Shoring supports, etc.)
Site & Structure Inspection	▪ Level surface or platform free from obstacles. ▪ Visibility between equipment and operator.
Pipe Fab. & Inst.	▪ Accessibility of welders by leaving two feet between connections and joints.
Structural Steel Fab. & Inst.	▪ Use of Object Based Design process so that design documents can be used to detail and fabricate material.
Material Tracking	▪ Specifications for use of barcode labels. ▪ Industry Standardization of technology. ▪ Design with more interchangeable sequences of work. ▪ Specifications for required participation.

SESSION 4

DESIGN AND DEVELOPMENT OF CONSTRUCTION ROBOTS

A TOOL TO IMPROVE EFFICIENCY IN LARGE SCALE MANUFACTURING

Roger Bostelman, Will Shackleford, Fred Proctor and James Albus

Intelligent Systems Division
Manufacturing Engineering Laboratory

Alan Lytle

Materials and Construction Research Division
Building and Fire Research Laboratory

National Institute of Standards and Technology[1]
100 Bureau Drive, MS-8230
Gaithersburg, Maryland 20899-8230
roger.bostelman@nist.gov, 301-975-3426

ABSTRACT: NIST is working directly with industry to improve repair and conversion operations of ships in dry dock. This work allows transfer of technology to construction and other industries requiring worker-access to large, external surfaces with minimum footprint and maximum system rigidity and control, while augmenting conventional suspended-scaffold systems and moving toward more autonomous large-scale manufacturing applications such as building construction.

KEYWORDS: worker access, ship repair, construction, robotics, cable controlled, large-scale manufacturing

1. INTRODUCTION

The Manufacturing Engineering Laboratory of the National Institute of Standards and Technology (NIST) has teamed with Atlantic Marine, Inc. in Mobile, Alabama to study efficient methods to repair ships in dry dock or along a pier. This project, called Knowledge-based Modular Repair [1, 2] is under the auspices of the Navy National Shipbuilding Research Program Advanced Shipbuilding Enterprise Initiative, where worker-, equipment-, and material access to external ship surfaces was determined to be a key focus area. The concept developed in this project is called the "Flying Carpet" and combines two main technologies: the NIST RoboCrane [3] and commercially available suspended scaffolding to produce an effective concept for worker access to ships, submarines, buildings, and other large objects.

The NIST Intelligent Systems Division developed the RoboCrane cable-controlled manipulator over several years [3, 4, 5, 6], during a project for the Defense Advanced Research Project Agency (DARPA) that studied crane suspended load control. Since the DARPA project, NIST has expanded RoboCrane technology into a viable solution to address large-scale manufacturing and many other challenges [7]. The RoboCrane applies the Stewart-platform parallel-link manipulator technology to a reconfigurable, cable-driven

[1] No approval or endorsement of any commercial product by the National Institute of Standards and Technology is intended or implied. Certain commercial equipment, instruments, or materials are identified in this report in order to facilitate understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

system. While RoboCrane can lift large, heavy and awkward loads, its stability and maneuverability allow advanced programming techniques more analogous to robots than cranes. The RoboCrane combines sensors, computers, and lightweight tensioned cable machines for performing heavy manufacturing and construction tasks. These tasks can include lifting and positioning heavy loads; manipulation of workers, tools and parts for improved worker accessibility and assembly; and fixturing, welding, cutting, grinding, machining, surface finishing and inspection.

Recent research has yielded the Flying Carpet concept as a movable scaffolding and worker positioning system that enables workers to maneuver themselves, parts, and tools throughout a large work volume for tasks such as ship repair and aircraft paint removal with up to 20-times improved efficiency over hand-built scaffolding. The Flying Carpet is a cable-supported platform that uses single-axis jog-, velocity- and force-control modes. A photograph of the 1:120 scale concept model is shown in Figure 1.

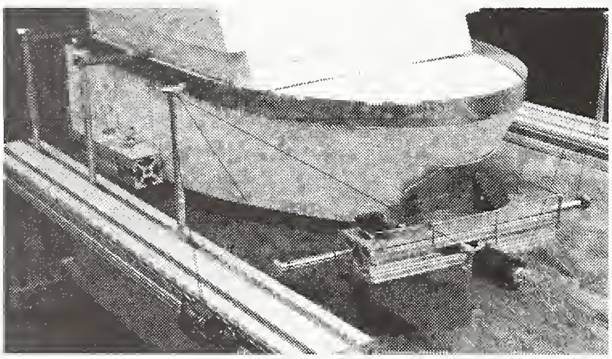


Figure 1. NIST Flying Carpet 1:120 scale concept model showing the ship bow/stern access configuration (right) and the ship side access configuration (left).

Beyond this step is envisioned a combined closed-loop system, whereby the Flying Carpet can autonomously and rapidly lift, position, and fixture heavy, bulky steel plates onto ships in dry dock during repair and conversion operations. Similarly, the system can be used for more autonomous assembly applications on construction sites. This paper will detail the Flying Carpet concept, status, and provide a look to the future toward the construction of buildings and more autonomous manufacturing.

2. CONCEPT

Small and full scale static physical models, a computer model for studying system work volume, and a full-scale working prototype were built to demonstrate the advanced functionality of the Flying Carpet as a tool for ship repair and other uses. Figure 2 shows a photograph of the full-scale working prototype Flying Carpet configured for ship bow/stern access. Its basic geometry includes four upper support points, instead of three, to match the dry dock configuration. These are connected to three work-platform points with six cables.

The four upper support points can be attached to towers mounted to a dry dock, on the ground, along a pier, to a gantry, or to the ceiling, walls, or other superstructures. Two front cable pairs provide platform lift while two rear cables mount lower to pull back on the platform, creating a rigid system. Cables can be multi-part lines for added safety and lift capacity. By suspending the platform from above, the RoboCrane improves operating efficiency by "flying" over ground-clutter or landscaping that typically hinders wheeled vehicles at the work site.

Hoists that control each cable's length can be mounted on the support structure or the work platform. The total hoist rigging capacity of the prototype, which uses two-part wire ropes, totals 8200 kg. In the prototyped configuration, the Flying Carpet carries its hoist motors, which provide a stabilizing counterweight. Motor location on the platform also eliminates the problem of mounting motors on the surrounding support structure, making the platform easier to reconfigure.

Figure 2. Full-scale Flying Carpet prototype top view shown in the ship bow/stern access configuration.

Welders, paint sprayers, or other equipment can be mounted to the Flying Carpet. The platform allows rapid fixturing of tools, equipment, or cargo to provide direct worker access to the equipment as needed at the site. The Flying Carpet cable configuration provides a constrained and easily maneuvered work platform as compared to conventional worker-access systems typically used for ship repair, thereby aiding ship welding and inspection tasks.

The platform provides minimized sway and rotation, and can exert forces and torques with full six degree-of-freedom control. Flying Carpet motion types include Cartesian and joint modes. Cartesian control allows the worker to very simply move the platform front-to-back, side-to-side, and up-and-down, as well as yaw about the vertical axis, all while maintaining level. Joint mode allows single-hoist motion for setup or cable replacement for normal maintenance.

Platform levelness is ensured by both the platform kinematic control and through a redundant level sensor. Operator control is through the tethered joystick, either worn by the operator or mounted to the platform. With the platform, an on-board or remotely-located operator can manipulate and hold attached materials, such as heavy steel plates, or tools, such as welders, grinders, robots and other cargo only dependent upon the platform rated capacity. Tension sensors in-line with each cable prevent hoist or platform overloading from occurring. An on-board supply hoist can also be attached to the platform to bring tools, workers, and supplies up to the work site while the platform is parked in position.

The full-scale prototype configured for ship bow and stern access measures 14.5 m wide x 7 m deep x 2 m high. Six hoists, each rated to lift 680 kg, can carry 680 kg of workers, materials, and equipment in addition to the 1400 kg weight of the platform itself, with at least a 5 times safety factor.

Performance measurements and cable configurations were tested on the full-scale testbed prior to planned testing in a shipyard dry dock. Constrained by the NIST facility spacing of 18 m simulated tower height x 21 m width x 14 m depth for front-to-back attachment point distance, the full-scale Flying Carpet prototype demonstrated: 10 m lift, 9 m forward-to-back motion, 5.5 m side-to-side motion, and yaw of more than ± 25° before the cables went slack. The translational work volume should scale well to the larger dry dock environment.

3. PLATFORM RECONFIGURATION

The Flying Carpet can be reconfigured from ship bow and stern access to a thin, ship side-access configuration as shown in Figure 3.

Reconfiguration from the bow and stern access to the side access configuration includes: removing the hoist platform (computer, power, and attached cables); removing the rear truss assembly; moving the hoists and pulleys to match the new configuration; re-spooling the cables; and pushing a button on the joystick to tell the computer that the configuration has changed. In this configuration, the 23 m depth is reduced to 2 m, allowing it to fit between the dry dock wing wall and the ship side.

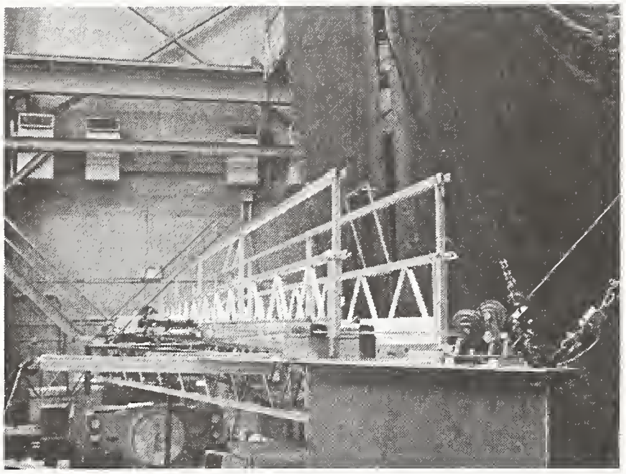


Figure 3. Full-scale Flying Carpet prototype top view shown in the ship side access configuration. The full-scale platform is near a mock ship (photo right) and in this test there was no dry dock wing wall to push against with the platform outriggers.

It took a team of 3 workers 13 man-hours to perform platform reconfiguration over about 4 clock hours. It is estimated that this time could be reduced to 3 to 6 man-hours (or 1 to 2 clock hours) with further experience. If a second platform were used for ship side-access along with a ship bow and stern access platform, this reconfiguration time would not be necessary.

The demonstrated work volume in the side access configuration with support points spaced at approximately 8 m x 21 m x 8.5 m high allows 6 m platform motion forward-and-back, 5.5 m side-to-side, and moves 6 m high. Yaw motion is limited to approximately 5° due to the reduced front-back depth and reduced rear platform depth. The platform in this configuration includes similar rigidity characteristics as in the bow and stern access configuration.

For platform heights above the dry dock wing wall, the cables are at smaller angles with the horizontal axis and therefore provide the lateral stiffness necessary for upper ship side-access, at a cost of increased cable tension. Vertical stiffness is reduced. In this case, all cables are mounted at the same height, where two cables attach to one of two towers as shown in the model in Figure 4.

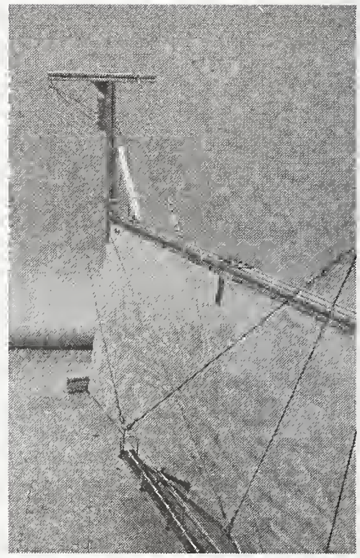


Figure 4. Photograph of the Flying Carpet 1:12 scale model shown in the ship side access configuration. The 1:12 scale model shows a tower attached to the dry dock wing wall and the platform pushing against the wall.

Two front cables can instead be crossed for even more rigidity, and towers (support points) can be separated by 30 m or more providing a very large range of motion side-to-side. Along the wing wall, similar platform rigidity can be accomplished by pushing against the dry dock wall with outriggers that could be adjusted manually or automatically.

4. APPLICATION TO THE CONSTRUCTION INDUSTRY

The concept of a Flying Carpet can also be applied to the construction industry. Figures 5 and 6 show a graphic of this concept and photographs of a scale model, respectively. For quick or prolonged and non-intrusive access to external building surfaces, the Flying Carpet can be attached to inflatable columns or other superstructures (e.g., building corners, aluminum or steel columns) and maneuvered about the building surface. Columns instead of building corners allow the Flying Carpet to be attached to unfinished structures. The columns can be made of a lightweight material, such as sailcloth, light enough for a single worker to set up from the ground. For example, a column 1 m in diameter, inflated to 70 kPa (10 psi), can support more than 5000 kg of weight. Given the stiffness of sailcloth, 1-m diameter columns up to 25 m long should have no tendency to buckle. Longer columns should be larger in diameter to preclude buckling. A small air pump could be used to fill the structure and lift the support points high above the building structure. Guy wires could be used to stabilize the structure. With this concept, a worker could theoretically reach all external building surface points for performing window or wall section insertion, window washing, brick laying, cleaning, or many other tasks associated with building construction or maintenance.

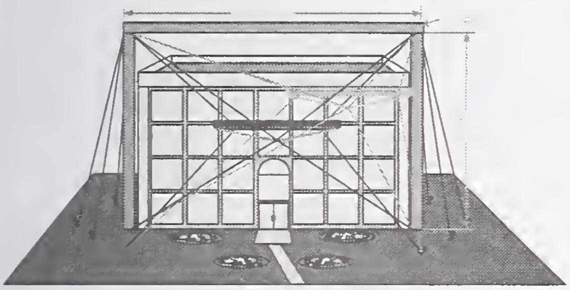


Figure 5. Graphic of a model Flying Carpet applied to external building surface access.

Figure 6. Model Flying Carpet front (top) and edge (bottom) applied to external building surface access. The wooden supports simulate inflatable columns or other superstructure that supports the Flying Carpet.

5. FUTURE CONCEPTS

One of the next steps in the Flying Carpet development effort is to provide navigation capability for platform operation in semi- or fully-autonomous modes. Although relative platform movement can be derived from winch encoders, encoders alone cannot reveal changes in platform position due to load variation or sway, nor will they easily map to the complex shape of a ship or submarine hull. A position reference system (absolute or relative) is desired to enable flight path trajectory planning for the Flying Carpet.

The Construction Metrology and Automation Group at NIST is currently instrumenting a RoboCrane platform with a three-dimensional, laser-based site measurement system (SMS) for absolute position control in all six degrees-of-freedom. [8] The project, part of the Automated Steel Construction Testbed, will be used to demonstrate autonomous steel pick and place operations. Follow-on experiments will incorporate registered LADAR (laser detection and ranging) scans of the work site for task analysis and navigation planning.

A similar navigation package will be employed for the Flying Carpet using the registered LADAR scans to map the hull as a boundary surface. The SMS will then be used to track the Flying Carpet along that surface. An advantage of this method is that the LADAR data from the scan also provides detailed 3D information that can be used to map damaged areas, create cutting/rolling templates for repair material, analyze surface imperfections, and generate as-built data for pre- and post-repair.

6. SUMMARY AND CONCLUSION

The Flying Carpet is a reconfigurable cable-controlled platform based on the Stewart Platform parallel mechanism. The Flying Carpet provides the dexterity, precision, and large work-volume needed for dry dock and/or pier side ship repair, as well as for other large-scale manufacturing applications. The Flying Carpet can be reconfigured and can attach to towers or existing superstructures to eliminate unnecessary equipment costs. Tools and equipment can be attached to the Flying Carpet quickly and easily for many worker-assisted tasks. The Flying Carpet operator can be located at the work site or at a remote location to provide safe and efficient worker placement. The Flying Carpet is a demonstrated technology, ready for commercialization. Advanced concepts toward autonomous construction are also being considered.

7. REFERENCES

1. Stieren, D., Caskey, G., McLean, C., and Neyhart, T. "Knowledge-Based Modular Repair: Advanced Technology Applications for Ship Repair and Conversion," Proc. of the 2000 Ship Production Symposium, Williamsburg, Virginia, August 24-25, 2000.

2. Stieren, D., Sovilla, L., "Rethinking Ship Repair: Knowledge-Based Modular Repair" Presentation, Proc. of the Shipbuilding Decisions 2001 Ninth Annual Commercial Shipbuilding Conference, Washington, D.C., December 4-5, 2001.

3. Stewart, D., "A Platform with Six Degrees of Freedom," Proc. of the Inst. of Mechanical Engineering, Volume 180(15), Part I:371-386, 1965-1966.

4. Albus, J. S., Bostelman, R. V., Dagalakis, N. G., "The NIST ROBOCRANE, A Robot Crane", Journal of Robotic Systems, July 1992.

5. Bostelman, R., Albus, J., Dagalakis, N., Jacoff, A., "RoboCrane Project: An Advanced Concept for Large Scale Manufacturing," Association for Unmanned Vehicles Systems International Proc., Orlando, FL, July 1996.

6. Roger Bostelman, James Albus, Bill Stone, "Toward Next-Generation Construction Machines," Proc. Of American Nuclear Society 9th International Topical Meeting on Robotics and Remote Systems, Seattle, WA, March 4-8, 2001.

7. Bostelman, R., Albus, J., Dagalakis, N., Jacoff, A., Gross, J., "Applications of the NIST RoboCrane," Proc. of the 5th International Symposium on Robotics and Manufacturing, Maui, HI, August 14-18, 1994.

8. Alan Lytle, Kamel Saidi, and William Stone, "Development of a Robotic Structural Steel Placement System," 19th International Symposium on Automation and Robotics in Construction, Washington, D.C., September23-25, 2002.

REMOTE CONTROL OF MACHINES FOR REMOVAL OF DAMAGES BEING RESULT OF DISASTERS, WARS, AND TERRORIST ATTACKS

Adam Bartnicki, Franciszek Kuczmarski, Tadeusz Przychodzień, Andrzej Typiak, Józef Wrona

Military University of Technology, Institute of Mechanical Engineering
KALISKIEGO STREET 2, 00-908 WARSAW, POLAND
phone +48 (0-22) 6839419, fax +48 (0-22) 6833211
E-mail: kuczmarski@wme.wat.waw.pl

Abstract: This paper presents an analysis of various natural and civilization disasters being angled as possibilities of usage of the remote control or autonomous machines. There are determined the requirements for such a systems of theses vehicles or machines which are able to remove effects of natural disasters and local military conflicts or terrorist attacks. Idea of remote controlled working machines with using of external control, information and observation system is also presented.

Keywords: Remote control, vision systems, machines, disasters

1. INTRODUCTION:

Dangerous events that occur more and more often in contemporary world, being result both natural disasters as well as terrorist attacks (e.g., contaminations, destructions and mining of terrains) and local armed conflicts extort necessity of wide utilization of different machines to remove (or limit) the negative results of them. However, the enormous majority of working machines (excavators, bulldozers, and mobile cranes) or combat engineering machines are used for such tasks being controlled directly by operator (without remote control or computer-aided systems). It is possible to expect that there will occur threats for operators and people who cooperate with them. Hence, we have to very intensely develop the remote control systems for such machines.

On the other hand, implementation the fixed elements on the board of these machines that adapt their to remote or automatic control is not often rational (reasonable) from economic point of view. Hence, taking the attempt to define the general structure of system to adapt these machines for remote control, but only when it is needed, very easily, with limited number of elements installed on the board of machine seems to be reasonable.

It is clear that the core of such a system (transmitter and control desk) and operator should be located out of machine (in safe zone for them).

It also requires to make the suitable outside subsystem of visualization of position of machine and its working equipment, to define their coordinates in coordinate system and to counteract in case of such critical incidents as damage or standstill of machine or troubles with communication between machine and operator (break of transmission of control signals).

The basic efforts of designers of control system of such machines should be focused on minimization of their onboard structure through:

- applying the simple and reliable elements of execution system;
- minimization of range of information transmitted from machine (decreasing the number of sensors and converters);
- introduction the "reliable core" of control system – operating in emergency states;
- possibilities of automatic reconfiguration of onboard system in case of breakdowns and faults.

This extorts necessity of extension of external structure of control system (placed beyond range of direct influence of surrounding of machine (vehicle), what causes:

- necessity of increasing of numbers of transmission channels and speed of sending of information and commands;
- growth of probability of occuring of errors - it requires to amplify signal power for keeping required range of transmission link;
- necessity of elaboration of detailed and complex steering procedures with many variants of operation.

Such designed control system of umanned vehicles and machines with various extent of automatic and autonomous operation can execute wide range of works such as removal of effects of natural disasters and local military conflicts or terrorist attacks (see table 1).

2. IDEA OF REMOTE CONTROL WITH USAGE OF EXTERNAL CONTROL AND OBSERVATION SYSTEM:

The requirements and conditions which are presented above have allowed to elaborate the idea of remote control of machines with usage of external control and observation system that can be used to remove effects of disasters, wars and terrorist attacks. It has been presented in figure 1. The area of operation has been divided into three zones: basic, supply and expected. In the basic zone a control post has been located. It has been equipped with remote control system to control of machines (1) and (2) and external TV observation cameras with geodesic measurement system of their position, orientation and observation. One of this TV camera is located in supply zone and designated to aid the operation of remained two TV cameras. Both control and vision signals are transmitted via transmission channels. It allows to obtain data concerning position and to control the execution elements e.g. levers, pedals of machines being controlled (1). In adverse conditions remote controlled unmanned vehicles for added observation of machine that is remote controlled in the expected zone can be used. This unmanned, mobile observation vehicle (2) plays the role of relay station for control and information signals between the controlled machine (1) and control post (3).

Standard vision systems perceive the word in two dimensions (because of lack of direct ability to measure the distance to the objects and their size). Thus nowadays the main research problem in vision system is three-dimensional perception of the picture (particularly its depth). The data of the three-dimensional location of the objects which are not included in two-dimensional picture are included e.g. in stereoscopic pictures.

Stereovision allows to achieve relatively high spatial vision and in comparison with one eye techniques (one-camera) it leads to quantity measurements, more direct and unambiguous. Despite of radar and laser distance measurements it is suitable for many applications. It particularly concerns to reconnaissance, supervision and objects' manipulation in an unknown area detection of vehicles' motion in unknown surrounding, including metric measurement in photogrammetry, robotics and vehicle remote control.

New vision system was elaborated for these needs. The system consists of two basic blocs: observation set and operator position (see figure 2).

Observation set serves to observe the terrain or chosen objects with CCD cameras. Its construction allows to observe chosen objects with both cameras (carrying out stereovision process) or observation the objects located in different directions. At the operator position there are TV cameras, operator panel and microprocessor controller VME.

In the figure 3 the developed conceptual scheme of elaborated vision system is presented. In the system three channels radio link has been applied. Two one direction channels serves to send television signal while two direction channel serves to send steering and measuring signals.

Observation set consists of a steering camera-head mounted symmetrically on a rotary beam. The change of the axis of beam rotation is done by its steering head. GPS system mounted to it reads its azimuth and specifies: set location in geographical coordinates, sea level and terrain slope. The data from GPS might be plotted on digital map which would allow to draw e.g. the route of a vehicle drive (with mounted vision system).

Operator stand allows to control the position of observation cameras, the change of their focal length and focus regulation. Steering might be carried out automatically – with usage of microprocessor controller or in manual way – through operator panel. The data about camera parameters and the location and orientation of observation set are sent to controller and moreover might by shown (in any configuration) on the monitor screen.

Microprocessor controller VME has been designated to automatic control of cameras' position and it is an element supporting the operator. Moreover it enables to process vision signals and determine on the basis of these signals location of observed object in relation to geodesic point of reference. The control procedures of steering cameras' position is used for automatic monitoring the observed object.

3. CONCLUSIONS:

Resulting conclusions from the above schemes and examples indicate that presented idea of remote control vehicles for removal effects of disasters, wars and terrorist attacks is proper. There were underlined the main problems during realization proposed control system i.e. elaboration of visual system for visualization of surrounding of vehicle or its working area. It makes possible the control of such machines in real time by average educated operator.

Presented solutions indicate there is necessary individual approach for designing of structures of control systems in depend on expected technological tasks.

The existing dangerous for human life that have been determined in this paper extort the needs to elaborate the remote and automatic systems for universally used machines and vehicles to make them more safe for operators during works after different kinds of natural and civilization disasters.

Presented vision system can be used for:

- localization of observed objects;
- surrounding observation and detection terrain obstacles;
- observation position and orientation of working equipment;
- visualization of working equipment position.

The additional advantage is that it can be mounted on both engineering machines and separately placed in the terrain.

Its main aim is supporting the operator while fighting or while earthmoving works in inaccessible or hazardous terrain. The assumption of the project was to allow the operator to see working machine (to have with it an eye contact).

On the basis of the picture obtained from cameras both can be specified – vehicle's position and external objects location.

The system may be used in armed forces to locate and indicate the targets as well as to support the fire control at the battlefield.

In further works on the presented system it is assumed that the spatial image will be obtained on the base of data from the observation cameras. It will allow to implement the pictures of observed pictures on TV monitors. Thus the soldiers and the operators of the remote controlled machines and vehicles will be able to perform their tasks more precisely.

4. REFERENCES:

1. Bartha M.; Eibert M.; Schaefer Ch.H.: *Robot-motion in unstructured environment*, Twelfth International Symposium on Automation and Robotics in Construction, Warszawa 1995.

2. Bartnicki A., Konopka S., Kuczmarski F., Typiak A.: *Badania doświadczalne sterowania koparką jednonaczyniową z systemem wizyjnym*, XV Konferencja Problemy Rozwoju Maszyn Roboczych, Zakopane 2002.

3. Masłowski A.: *Mobilne roboty interwencyjno-inspekcyjne*, Konferencja Automation 97, Warszawa 1997.

4. Bartnicki A., Konopka S., Kuczmarski F., Typiak A.: *Remote control of working machines with vision system*, 18 th ISARC International Symposium on Automation and Robotics in Construction, Kraków 2001.

5. Roberts J.M.; Winstanley G.J.; Corke P.I.: *Issues in obstacle detection for autonomous mining and construction vehicles*, 16 th IAARC (IFAC) IEEE International Symposium on Automation and Robotics in Construction, Madrid 1999.

Table 1. Classification of disasters being angled as possibilities of usage of the remote control or autonomous machines*

DISASTERS	
NATURAL	CIVILAZATION
Landslips, earthquakes, volcanic eruptions	Contaminations - radioactive, - chemical, - biological
Floods	Calamities and destructions of objects: - buildings, -communications
Deep snowfalls, snowstorms, blizzards	Fires
Strong winds and hurricanes	Terrorist attacks and different threats of public safety
Long-lasting freeze or droughts, hailstorms or glazes	Destruction damages of infrastructure in result of terrorist attacks and local military conflicts
Plague of rodents and insects	

*) intensity of gray color indicates extent of usage of the remote control or autonomous machines

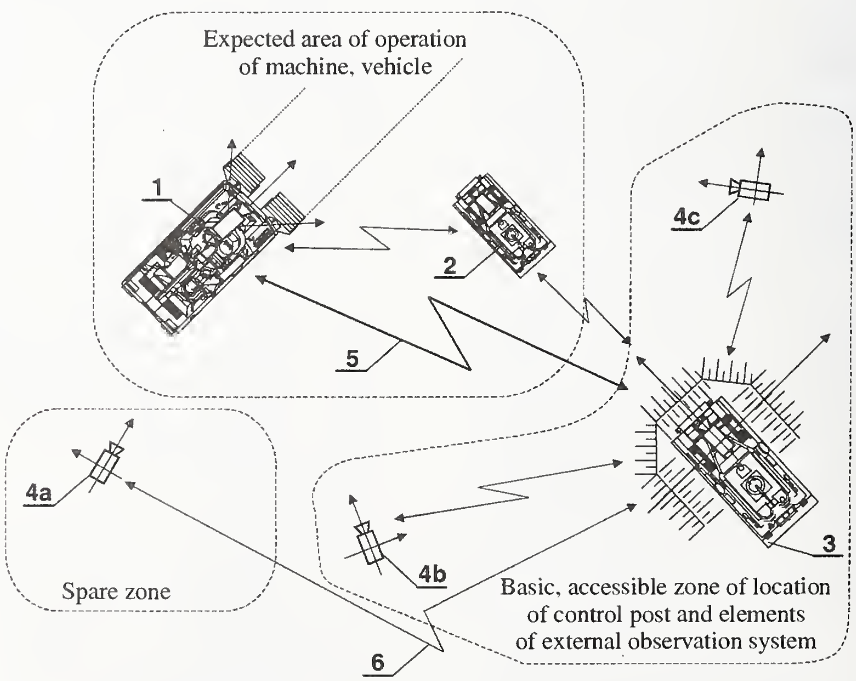


Figure 1. Idea of remote controlled working machines with usage of external control and observation system

1 - machine being controlled; 2 - umanned, mobile observation vehicle 3 - basic control post (operator equipped with control panel or special vehicle); 4a, b, c - external remote controlled observation cameras with geodesic measurement system of their position and orientation; 5, 6 - channels for controlling of the cameras and for transmission of vision and position data

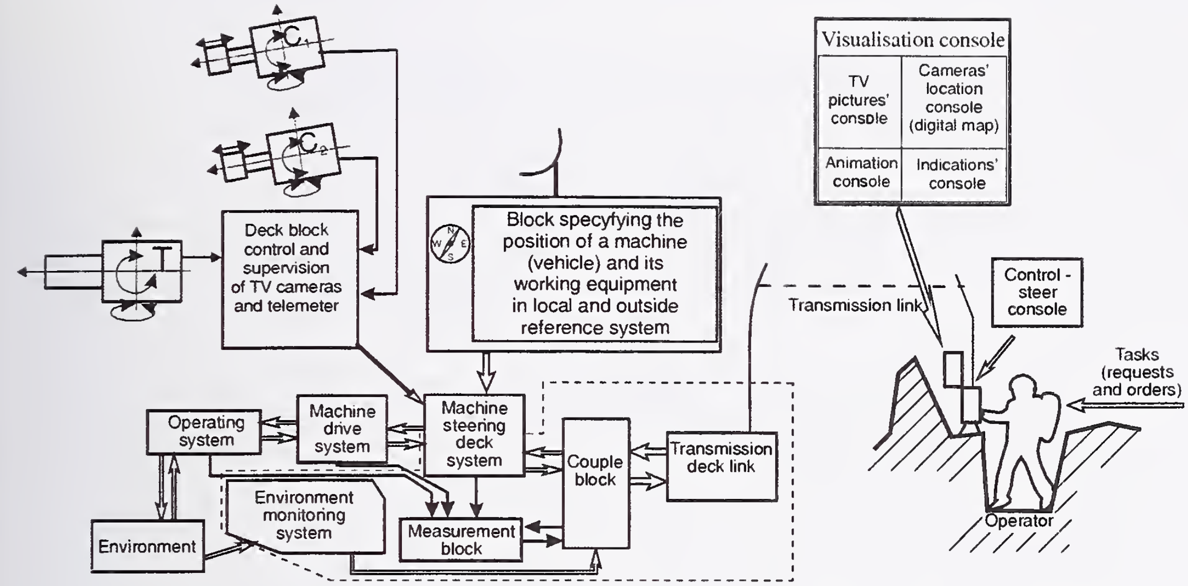


Figure 2. Structure of remote control of mobile working machines' scheme with vision system

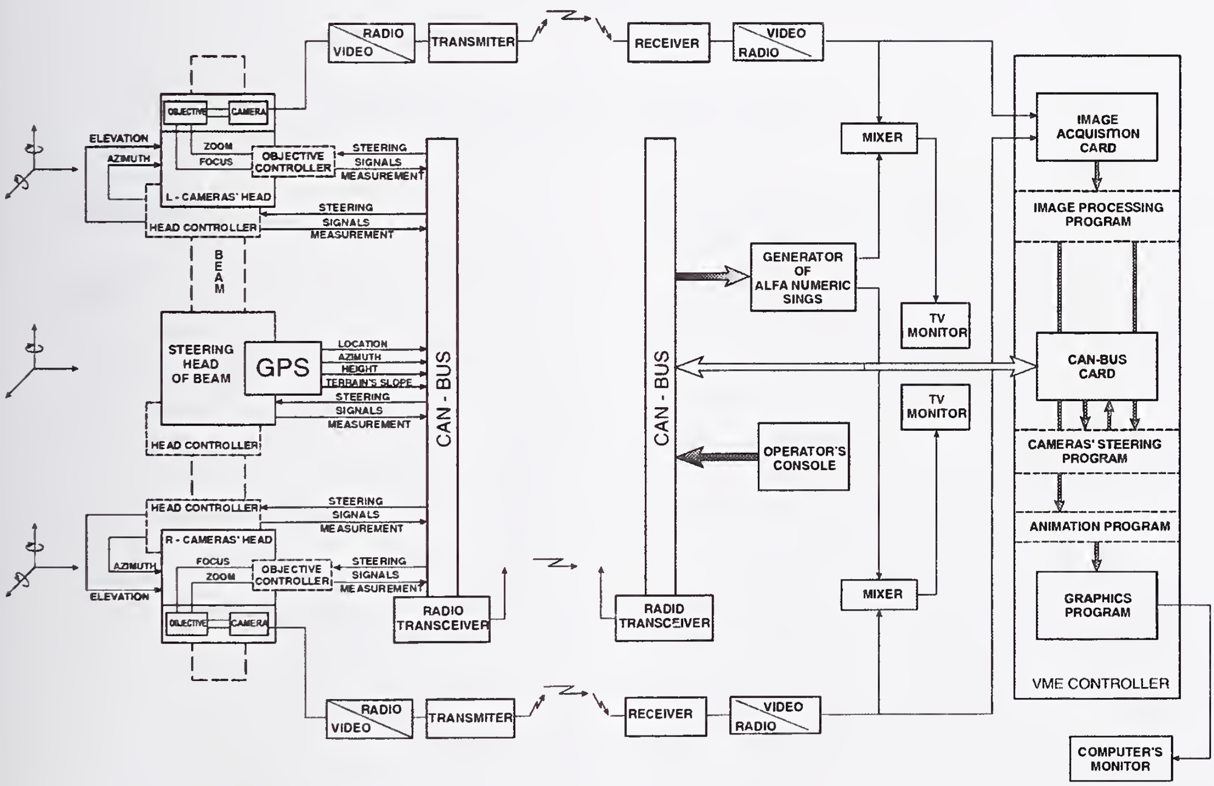


Figure 3. Scheme of vision system

ADVANCEMENTS IN TELE-ROBOTIC PIPELAYING

Leonhard E. Bernold,
Associate Professor
Bin Li, *Intern. Scholar*
Construction Automation Robotics Laboratory (CARL.ce.ncsu.edu),
North Carolina State University, Raleigh, NC 27511

Abstract: The labor intensive construction industry has a very high accident rate. One of the key reasons for this is the exposure of workers to a hazardous environment such as heights, confined narrow spaces, and exposure to health threatening fumes, dust, and noise. Tele-robotic operation, which allows an operator to control a mechanical tool from a safe distance, provides a technical alternative. Because of the need to be competitive on every project-bid, contractors have to be assured that new technologies not only work in the rugged environment of a construction site but that they also reduce cost. This paper will present an example of such a technology, a pipe manipulator that allows the remote installation of large concrete pipes in deep trenches. A second prototype has been built and successfully field-tested. The comparative evaluation of the new technology shows that it is not only safer but also more economical to use. With the goal to further improve agility, flexibility, and ease of operation, a third prototype has been readied for further field tests

Keywords: tele-robotics, safety, trenching, remote control, pipe-laying, trench collapse

Introduction

According to the Bureau of Labor Statistics (1999), in 1997 there were 1,137 deaths in the construction industry. This number equals 18% of the total fatalities from all industries and indicates that construction has the highest number of fatalities. Another source shows that the construction industry employed approximately 5% of the industrial work force, but generally accounted for nearly 20% of all accidental deaths (National Safety Council, 1997). An analysis by the National Institute for Occupational Safety and Health (NIOSH) of workers' compensation claims for 1976-1981 shows that excavation cave-ins were the cause of about 1,000 work-related injuries each year. Of these injuries, 140 resulted in permanent disability and 75 in death.

The traditional ways to prevent collapse and make working in a trench safe are (1) providing physical supports for each side of the trench using shoring or a trench box, or (2) sloping the sides to a safe angle.

OSHA's excavation standard requires employers to provide sloping (or benching), shoring, or shielding to protect employees in excavations 5 feet or more in depth. The only exception is for a trench dug in stable rock, where there is no loose soil or likelihood of a cave-in. Excavations less than 5 feet deep need not be protected unless a competent person has determined there is a cave-in hazard. As shown in Figure 1, depositing spoil right next to the trench make even a 5 feet deep trench a death trap. Because of the complex nature of digging soil, accidents still occur. One way to eliminate the risk, is to introduce a technical intervention such as teleoperated pipe manipulator able to lay pipe without a human in the trench. However, in order to be a viable economic alternative to the traditional methods, such technology needs to be cost effective as well.

Prototyping a Technical Intervention

A first device able to handle large concrete pipes remotely was designed and fabricated in 1994. This first-generation pipe manipulator (PipeMan), shown in Figure 2, was subsequently improved by adding a laser and video system in order for an operator to control the entire device remotely (Lee at al., 1999). Figure 3 presents the overall architecture of the 2 generation PipeMan.

Mechanical Components of PipeMan

The actuation system consists of the following five functions:

1) *H. Actuator (A_Rot1)*: A hydraulic actuator provides the limited ±100 degrees rotation of the pipe in order to line up the pipe.
2) *H. Actuator (A_Rot2)*: A hydraulic actuator provides a locking mechanism on the back of the manipulator to prevent it from slipping off the bucket.
3) *H. Cylinder (C_Trans)*: A hydraulic cylinder provides a linear activation of the pipe to joint the new pipe to the pipe already laid.
4) *H. Bladders (B_Clamp)*: Hydraulic bladders provide an inflating and deflating mechanism to clamp the manipulator to the bucket.
5) *E. Winch (W_Hold)*: An electric winch is used to attach the pipe with a quick release.

Man-Machine Interface

The man-machine interface includes manipulation and visualization functions. Two microvideo cameras allow the operator to have a real-time view during final alignment of the pipe joint. One video camera covers the front of the pipe for jointing and the other faces the laser targets to monitor the laser beam for alignment.

In order the provide the operator a convenient and safe method for operating the manipulator a separate control box, shown in Figure 4, was designed and fabricated. The control box included a DC/DC converter in order to distribute 24VDC to 5, 9, 12, and 24 VDC as outputs. The power source to be used is 24VDC from the excavator. A switchstick control is added for the remote operator's easy control of the manipulator. The control box also includes three three-way switches, and one main power on/off switch. Switchstick #1 controls rotational actuation of the hydraulic actuator (A_Rot1). Switchstick #2 is linked to the translational actuation of the hydraulic cylinder (C_Trans). Separate three-way switches control the actuation of the hydraulic cylinder (A_Rot2), clamping actuation of the bladders (B_Clamp), and locking/unlocking of backstop (A_Rot2). The LEDs (light-emitting diodes) are used as a fault detection system. As long as an electric signal or power is applied, the LED light remains on. The hydraulic fluid for the 4 different actuation units (two actuators, one cylinder, and two bladders) is provided by a hydraulic manifold mounted on to the PipeMan.

Comparative Field Tests

As mentioned earlier, only field tests with experienced laborers and operators will show if the innovations works effectively and economically. For this purpose a comparative field test was conducted on the job-site of East Park Industrial Subdivision in Raleigh, NC. The comparative experiment included two different ways of laying the pipes: 1) traditional method, and 2) manipulator utilization. Site conditions including soil, trench width and depth, pipes, crew members, weather were the same except for a way of performing pipe laying.

The crew members laid concrete pipes the way they normally do one day, and the manipulator was utilized to do the same tasks the next day. Figure 5 presents the two methods. The soil condition of the job site was sandy clay and the concrete pipes were 36-inch (0.9 m) diameter and 8 feet (2.4 m) long. Figure 5 b) depicts the operator lining up a new pipe by adjusting direction and grade of the pipe using the laser as a guide.

Qualitative Assessment

The technical performance of the new system was assessed in terms of: 1) Overall performance, 2) learning curve effect, 3) acceptance and adoption, and 4) technical problems.

Overall Performance

Using both, his own laser and the laser-video targets, the operator immediately felt comfortable with the necessary adjustments and the accuracy of the installed pipes. The availability of an image showing a close-up of the pipe joint was praised and felt to be a special asset when laying pipes into deeper trenches.

Learning Curve

The workers became familiar with the new process after only a few repetitions. They laid the last 2 pieces of pipes at a consistent shorter time than the first 3 cycles.

Acceptance and Adoption

The operator, pipe-layers, and helpers accepted the new technology whole-heartedly.

Technical Problems

Although a brand-new winch was used, the weight of the large pipe used by the contractor was too much for the its base-plate. The test made clear that having two winches might be necessary. Another issue was the size of the bucket which made the backstop unworkable. A chain provided the perfect alternative solution. The inflatable bags were also not performing as expected.

Cost Comparison

The main cost items consist of: a) excavation and bedding, b) pipe laying, c) pipe manipulator rental, d) backfill & compaction, and e) insurance. Volume of excavation for the traditional method, 21 m^3, includes the benches as shown in Figure 5 while the volume for the alternative method, 11.6 m^3, is based on vertical trench walls with a minimal width as required for proper haunching. As expected, the resulting cost saving, $ 7.52/m or 54%, is significant.
Overall, the total cost for the excavation and laying of 36 inch (0.9 m) concrete pipes into a 6 feet (1.8 m) deep trench during the field tests are $53.7 per meter ($16.37/ft) for the traditional method and $ 65.54 per meter ($20/ft) for the observed tele-robotic method. As was anticipated, the tele-robotic option turns out to be more expensive mostly as a result of the higher cycle time, the result of the broken winch. However, simple improvements should undoubtedly result in reducing the time for laying one pipe.

Third Generation PipeMan

Based on the encouraging results of the field test, it was felt that the basic concept of PipeMan was solid. One main problem, however, was the weaknesses of the winch. A second issue was the inability of adjusting the grade without the bucket. This capability would allow an excavator to be positioned on the side of the trench. The following section will present an alternative method of holding the pipe.

A Two Pronged Carriage

The flexibility of the system design allowed us to consider exchanging the carriage holding the winch while keeping the base, holding all the controls. It was decided to fabricate an alternative to the original carriage that was based on the flexibility provided by prongs. Figure 6 presents the new design. As indicated, the pipe will be lifted simply by inserting the prongs into the pipe. A mechanical stop-plate is automatically engaged when the weight of the pipe is taken over by the prongs. This stops the pipe from sliding off.

Adding a Third Degree of Freedom

One weakness of PipeMan was the fact that it could only be used in the axis of the trench since the slope of the pipe had be adjusted using the bucket. In order to allow a backhoe to lay pipes from the side of the trench, a degree of freedom had to be added. Figure 7 indicates that this was accomplished by allowing the prongs to rotate at one while actuating the other end with the help of a hydraulic cylinder.

Summary

Traditional trenching and pipe laying requires workers to enter the trench, resulting in many fatalities due to the nature of the soil in the trench walls and other work related circumstances. The telerobotic concept promises to drastically reduce the risk to human life by keeping the worker outside of the trench. This paper presented the major components and functions of the newest version of the teleoperated pipe manipulator that has been designed and fabricated to handle pipes that use o-ring connection. The new technology proved its technical feasibility by laying 8 piece of concrete pipes without any workers in the trench. The field test highlighted too that with the help of technical modifications, the system would also be cost effective. These modifications have been implemented to create the third generation PipeMan.

Acknowledgments

Funding for the work has been provided by Public Health Service-National Institutes of Health under contract 5 R01 CCR413051-02 and 1R01 0H04201-01. Its content is solely the responsibility of the writers and do not necessarily represent the official views of NIH.

References

Lee, J., Lorenc, S.J., and Bernold, L.E. (1999). "Saving Lives and Money with Robotic Trenching and Pipe Installation," ASCE Journal of Aerospace Engineering, Vol. 12, No. 2, ISSN 0893-1321, April 1999, pp. 43-49.

U.S. Department of Labor, Occupational Safety and Health Administration, Title 29 of the Code of Federal Regulations (CFR) Part 1926.650-652.

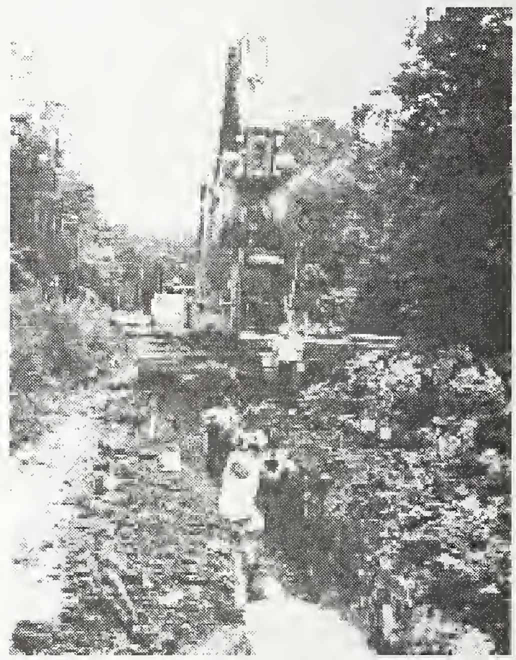


Fig 1. Dangerous Work Laying Pipes in an Open Trench

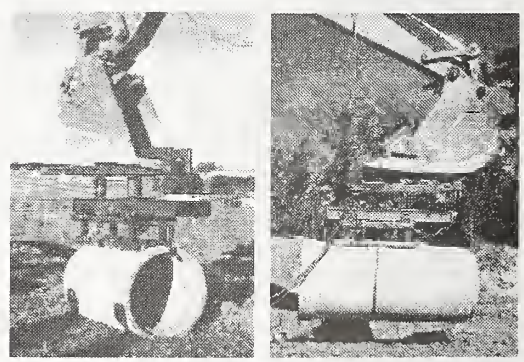


Fig 2. PipeMan Generations 1 and 2

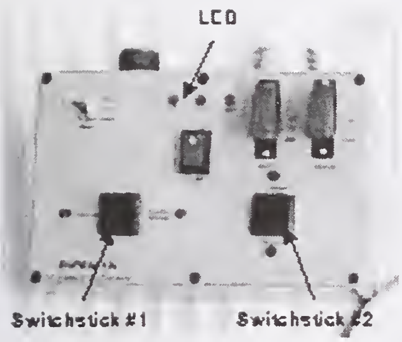
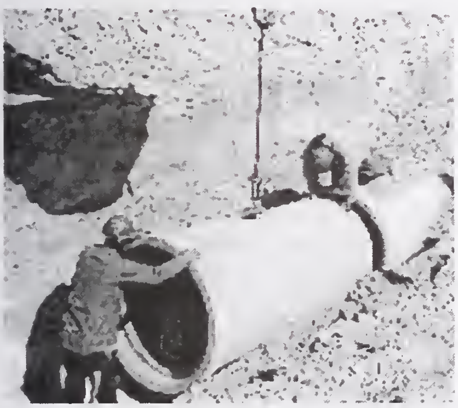
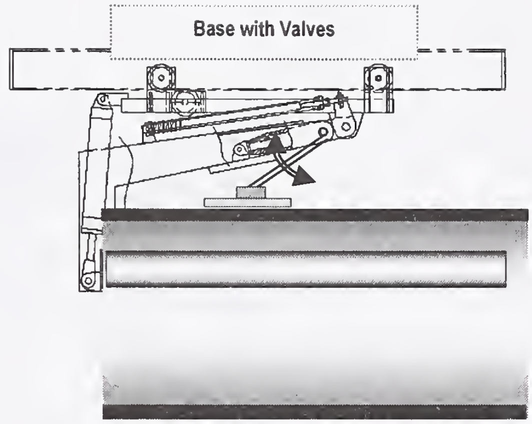


Fig. 4 Control Box for Operator

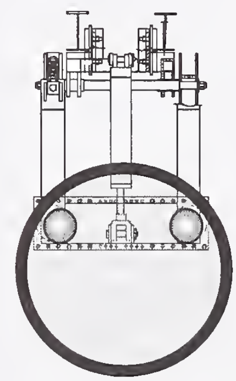


a) Laying Pipe Traditionally

b) Laying Pipe with PipeMan
c)
Fig 5. Comparative Field Test



a) Side View with Loaded Pipe



b) Front View with Loaded Pipe

Fig. 6. Pipe Carriage With Prongs

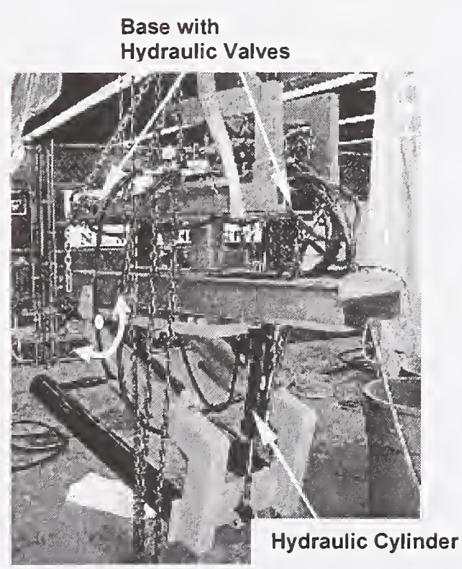


Fig 7. Third Degree of Freedom

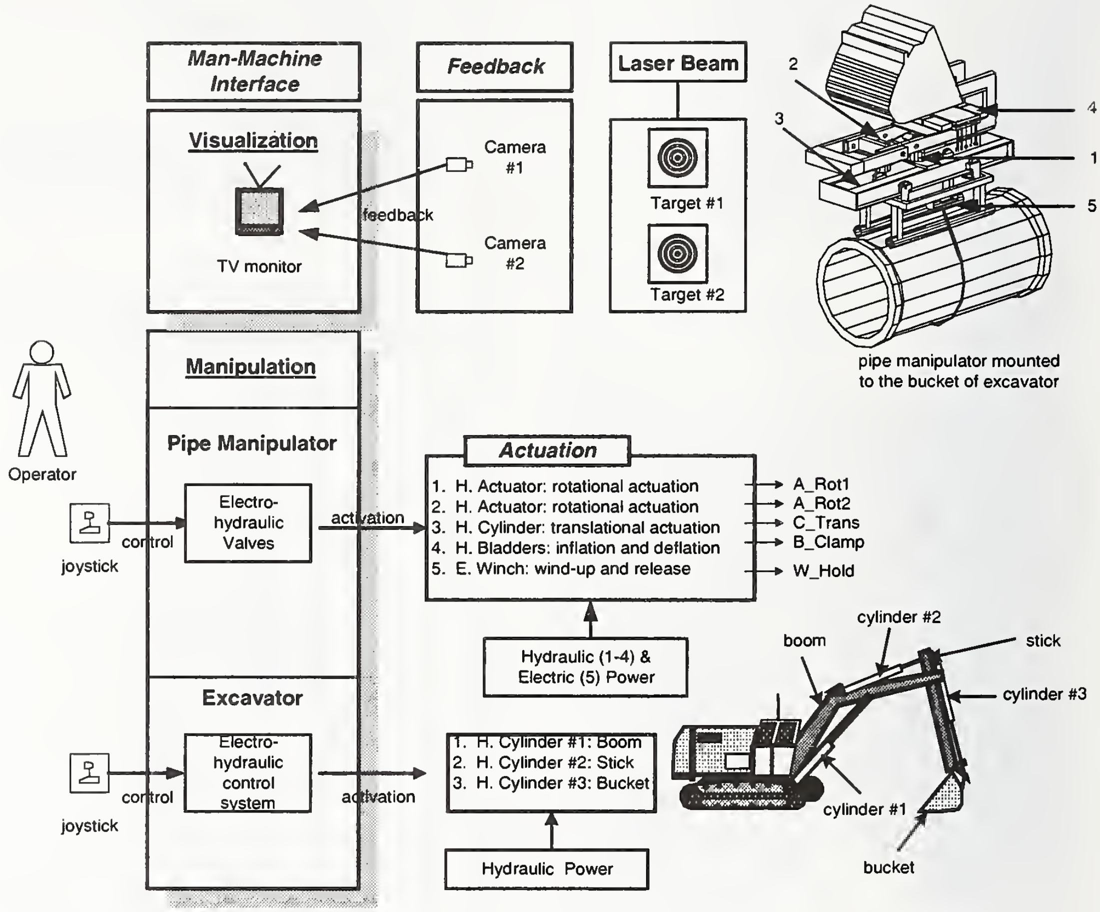


Fig 3. Layout of PipeMan Attachment

MOTION PLANNING TAKING INTO ACCOUNT THE DYNAMIC MODEL OF VEHICLES : APPLICATION TO THE COMPACTOR

Charles-Eric Lemaire * **Pierre-Olivier Vandanjon** *
Maxime Gautier ** **François Peyret** *

* *Laboratoire Central des Ponts et Chaussées (LCPC)*
Route de Bouaye, BP 4129, 44341 Bouguenais, France
charles-eric.lemaire/pierre-olivier.vandanjon@lcpc.fr
** *Institut de Recherche en Cybernétique et Communication de Nantes (IRCCyN)*
1 rue de la Noë, BP 92101, 44321 Nantes Cedex 3, France
maxime.gautier@irccyn.ec-nantes.fr

ABSTRACT: Motion planning is a necessary step before automation in road construction. This paper deals with motion planning for compactors. The compactor is an articulated frame steering vehicle used in road construction. Generally, motion planning uses the kinematic model of the robot. In the case of the compactor, this model is far away from reality because it does not take into account compactor masses and particular drum-soil contact forces. A motion planning which uses both the kinematic and the dynamic model of the compactor is presented in this paper. The advantages of this motion planning are:

- To generate feasible motion for the robot which implies a smoother control law for mechanics,
- Limited slips.

KEYWORDS: Compactor, Mobile robots, Motion planning, Dynamic model

1. INTRODUCTION

This article focuses on the compaction process during the road construction. This process is made up of several trajectories of several compactors. In the frame of automation in road construction, it can be seen that the compactor must follow an efficient trajectory defined in position, velocity and acceleration according to compactor degrees of freedom to guarantee the homogeneity and the expected density of considered compacted material. Those features should be improved taking into account a dynamic model compared with a simple path tracking using kinematic model. The compactor motion planning, presented here, is based on the dynamic modelling and identification of the compactor (Guillo *et al.*, 1999). Necessary elements of modelling are presented in this paper before the presentation of the method of motion planning with simulation results.

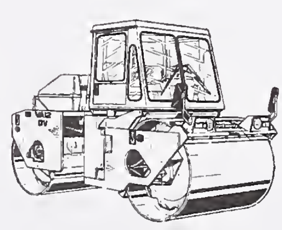


Fig. 1. A typical compactor: Albaret VA12 DV

2. COMPACTOR MODEL

A planar motion of the compactor is considered. Let R_g be a Galilean reference frame attached to the rolling plan Π. In these conditions, a vector of three generalized coordinates is needed to specify the compactor posture.

The front frame is chosen to be the reference body C_r of the compactor, *i.e.* its situation gives the compactor posture.

According to classical robot manipulator description (Canudas de Wit *et al.*, 1997), the compactor is considered as a mechanical system Σ composed of a tree structure of $n = 7$ rigid bodies C_j where C_0 is the base body and with the following body definitions (see Figure 2):

- C_1, C_2 are two virtual bodies (i.e. without mass and inertia) used to define the compactor position with respect to the frame R_0,
- C_3, C_5 are the front and rear drums,
- C_4, C_7 are the front and rear frames,
- C_6 is a virtual body (used to define a second frame R_6 attached to C_5).

The system Σ is provided with a frame R_j respectively attached to each of the $(n+1)$ bodies C_j. Let R_j be defined as $R_j = (O_j, x_j, y_j, z_j)$ (see Figure 2).

Classical tree structure description using the DHM notations (Khalil and Kleinfinger, 1986) applied to the system Σ defines the geometric parameters of the compactor (see Table 1 and Fig. 2) with respect to the position and orientation of the body C_0.

Table 1. Geometric parameters of the compactor

j	$i = a(j)$	σ_j	α_j	d_j	θ_j	r_j
1	0	1	$\pi/2$	0	$\pi/2$	r_1
2	1	1	$\pi/2$	0	$\pi/2$	r_2
3	2	0	$\pi/2$	0	θ_3	0
4	3	0	$-\pi/2$	0	θ_4	0
5	3	0	0	$-\mathrm{D}_5$	θ_5	0
6	5	2	0	$-\mathrm{D}_6$	0	0
7	6	0	$-\pi/2$	0	θ_7	0

According to the DHM description of the compactor, the vehicle motion is completely described by the vector q (see eq. 1) of six generalized coordinates where the vector ξ gives the compactor posture.

$$q = \begin{bmatrix} r_1 & r_2 & \theta_3 & \theta_4 & \theta_5 & \theta_7 \end{bmatrix}^{\mathrm{T}} \qquad (1)$$
$$\xi = \begin{bmatrix} r_1 & r_2 & \theta_3 \end{bmatrix}^{\mathrm{T}}$$

2.1 *Geometric model*

The Direct Geometric Model (DGM) gives bodies situation with respect to the reference frame R_g as a function of q, the vector of joint variables.

For example, equation 2 gives the situation of the rear frame (C_6) of the compactor with respect to the frame R_g. The symbolic modelling software SYMORO+ (see (Khalil and Creusot, 1997)) had been used to determine the expression of equation 2 from the table 1 of the geometric parameters of the compactor.

$$\begin{aligned} {}^g\mathrm{A}_6 &= \begin{bmatrix} \cos(\theta_3+\theta_5) & -\sin(\theta_3+\theta_5) & 0 \\ \sin(\theta_3+\theta_5) & \cos(\theta_3+\theta_5) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ {}^g\mathrm{P}_6 &= \begin{bmatrix} r_1 - \mathrm{D}_5\cdot\cos\theta_3 - \mathrm{D}_6\cdot\cos(\theta_3+\theta_5) \\ r_2 - \mathrm{D}_5\cdot\sin\theta_3 - \mathrm{D}_6\cdot\sin(\theta_3+\theta_5) \\ 0 \end{bmatrix} \end{aligned} \qquad (2)$$

2.2 *Kinematic model*

In the previous section, the geometric description of vehicles using robotics formulation is presented. From this description, the *Direct Kinematic Model* (DKM) of vehicles can be developed. It is composed of n_t relations (see eq. 3) corresponding to the velocity of each of the n_t vehicle terminal bodies (wheels, drums, tools,...).

The equation 3 gives the kinematic wrench of a vehicle terminal body C_j with respect to the reference frame R_g.

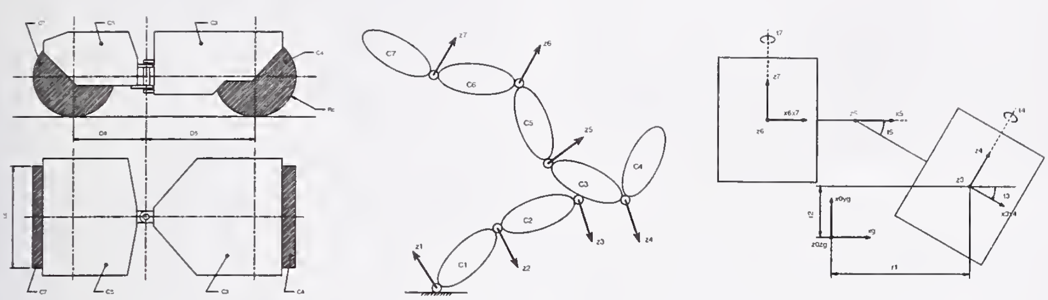
$$\begin{bmatrix} \mathrm{V}_{j,g}^{O_j} \\ \omega_{j,g} \end{bmatrix} = \Phi_j^{O_j}\,\dot{q} \qquad (3)$$

where $\mathrm{V}_{j,g}^{O_j}$ and $\omega_{j,g}$ respectively are the velocity of the point O_j and the rotation velocity of the body C_j with respect to the reference frame R_g. $\Phi_j^{O_j}$ is called *base jacobian matrix* of the vehicle.

2.2.1. *Conditions of pure rolling and non-slipping*

Let be assumed that the contact between the rigid drum and the soil is reduced to a single point B_j of the rolling plan Π (see Figure 3). The contact between the drum and the soil is supposed to satisfy both conditions of *pure rolling* and *non-slipping* along the motion. This means that the velocity of the contact point B_j is equal to zero and implies that the two components of this velocity, respectively parallel to the plan of the drum and orthogonal to this plan (see eq. 4), are equal to zero.

$$\begin{cases} \mathrm{V}_{j,g}^{B_j}\cdot x_i = 0 \\ \mathrm{V}_{j,g}^{B_j}\cdot y_i = 0 \end{cases} \qquad (4)$$



(a) Bodies definition (b) Equivalent tree structure (c) Projection frames definition

Fig. 2. Geometric description of the compactor using the MDH notations.

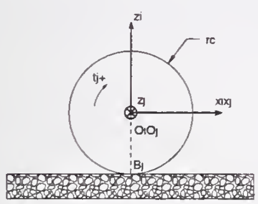


Fig. 3. Contact between a rigid drum and the soil reduced to a single point.

The constraints can be written in the general matrix form

$$J(q)\dot{q} = 0 \tag{5}$$

This means that whatever the type of vehicle, the velocity $\dot{q}$ is restricted to belong to a distribution Δ_c defined as

$$\dot{q} \in \Delta_c = \mathrm{span}\{\mathrm{col}(S(q))\} \tag{6}$$

where the columns of the matrix $S(q)$ form a basis of $\ker(K(q))$. This is equivalent to the following statement: for all time t, there exists a time-varying vector $\eta(t)$ such that

$$\dot{q} = S(q)\eta \tag{7}$$

The relation 7 is called the *Inverse Kinematic Model under constraints* of the vehicle.

2.3 *Dynamic model*

The *Inverse Dynamic Model* (IDM) of a vehicle is written as following

$$M(q)\ddot{q} + H(q,\dot{q}) = \mathbb{L} + \mathbb{Q}^c \tag{8}$$

where:

- $M(q)$ is the mass matrix of the system Σ.
- $H(q,\dot{q})$ is the vector of centrifugal, Coriolis and gravity terms.
- $\mathbb{L}$ is a vector depending on the internal forces between the vehicle bodies: motor torques, friction, lumped elasticity.
- $\mathbb{Q}^c$ is a vector depending on the external contact forces between the soil and the drums.

2.3.1. *Internal forces* The vector $\mathbb{L}$ of internal forces is composed of three components :

- the actuation vector $\mathbb{L}^a$ which is composed of motor torques,
- the friction vector $\mathbb{L}^f$ which is composed of friction component on different joints,
- the elastic forces vector $\mathbb{L}^e$ which represent the stiffness of joints.

2.4 *Contact forces between drums and the soil*

Using the principle of virtual powers, the vector $\mathbb{Q}^c$ of external forces is developed to obtain

$$\mathbb{Q}^c = \sum_j {}^i\Phi_j^{O_j T}\ {}^i\mathbb{T}^{O_j}_{\Pi_j \to C_j} \tag{9}$$

where ${}^i\mathbb{T}^{O_j}_{\Pi_j \to C_j}$ points out the wrench of the resultant contact forces between the soil Π_j on the drum C_j at point O_j projected in frame R_i.

2.5 *Linearity property of the inverse model*

The expression of kinematic and potential energies are linear in relation to a set of ($n_p = 11$) parameters, X_s. Consequently, the expression of the Inverse Dynamic Model is also linear in relation to the same set of parameters and then it is possible to write it as following

$$Y_s = D_s(q,\dot{q})X_s \tag{10}$$

with $Y_s = \mathbb{L}^a$.

Using this property of the Inverse Dynamic Model, a Weighted Least Squares method of identification is proposed by (Gautier, 1997) to obtain the values of the dynamic parameters X_s of robot manipulators and applied to a mobile machine, the compactor, in (Guillo *et al.*, 1999).

2.6 *General expressions of the dynamic model*

The general expression of the Inverse Dynamic Model is given by the relation 8. According to the expression of the internal and external forces which take part in the model, there are two possible expression of this model:

- The conditions of pure rolling and non-slipping are satisfied. Then, the Inverse Dynamic Model has the following expression

$$\begin{cases} M(q)\ddot{q} + H(q,\dot{q}) = \mathbb{L}^a + \mathbb{L}^f + \mathbb{L}^e + J(q)^T\lambda \\ J(q)\dot{q} = 0 \end{cases} \tag{11}$$

- The conditions of pure rolling and non-slipping are not satisfied. Then, the Inverse Dynamic Model has the following expression

$$\begin{cases} M(q)\ddot{q} + H(q,\dot{q}) = \mathbb{L}^a + \mathbb{L}^f + \mathbb{Q}^c \\ \mathbb{Q}^c = \sum_j {}^i\Phi_j^{O_j T}\; {}^i\mathbb{T}^{O_j}_{\Pi_j \to C_j} \end{cases} \tag{12}$$

3. MOTION PLANNING

Motion planning is to determine joint position, velocity and acceleration. The method developed is composed of four steps

(1) Path determination in order to respect the geometric constraints of the compactor task.
(2) Timing law determination in order to respect the kinematic constraints of the compactor task.
(3) Generalized coordinates determination using the inverse kinematic model under constraints.
(4) Actuator torques and drum-soil contact forces determination in order to compare them with physical limitations.

3.1 *Path determination*

According to the geometric model, a planar motion of the compactor is considered. The path is determined by the compactor posture

$$\xi = \begin{bmatrix} r_1 & r_2 & \theta_3 \end{bmatrix}^T = \begin{bmatrix} x(s) & y(s) & \theta(s) \end{bmatrix}^T \tag{13}$$

where :

- θ is the path orientation,
- x and y are the path cartesian coordinates.

The path is characterized by its curvature K along the arc length s

$$d\theta = K ds \tag{14}$$

$$\theta(s) = \theta_0 + \int_0^s K(u)du \tag{15}$$

$$x(s) = x_0 + \int_0^s \cos(\theta(u))du \tag{16}$$

$$y(s) = y_0 + \int_0^s \sin(\theta(u))du \tag{17}$$

Concerning asphalt compaction, the compactor has to compact the bituminous mix spread by the paver. But a paver is wider than a compactor, so the compactor path is composed of juxtaposed tracks (see figure 4).

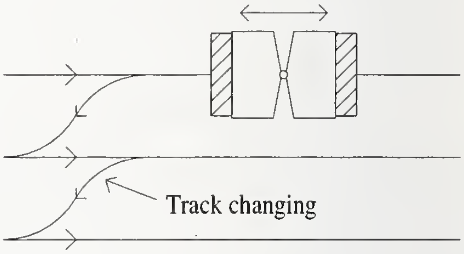


Fig. 4. Compactor typical path

The compactor path is essentially composed of straight lines and track changing. Determination of the path along a straight line is obvious.

$$K = 0 \tag{18}$$

$$\theta(s) = \theta_0 \tag{19}$$

$$x(s) = x_0 + s\cos(\theta) \tag{20}$$

$$y(s) = y_0 + s\sin(\theta) \tag{21}$$

A track changing path is characterized by its width D and its maximum curvature K_{max}. Parametric equations are determined in this way : there are two symmetrical parts in the path, if the path length is $4l$ then the curvature $K(u)$ is an odd function for $u = s - 2l$. These specifications are enough to compute $K(s)$ for $s \in [0, 2l]$.

The path curvature must be smooth to be compatible with the compactor dynamics. Constraints of symmetry lead to a second order function for the curvature such as :

$$K(s) = k_0 + k_1 s + k_2 s^2 \tag{22}$$

$$K(0) = 0,\ K(2l) = 0,\ K(l) = K_{max} \tag{23}$$

Solution for $K(s)$ with $s \in [0, 2l]$ is :

$$K(s) = 2K_{max}\frac{s}{l} - K_{max}\left(\frac{s}{l}\right)^2 \tag{24}$$

Cartesian coordinates could not be computed explicitly, so they are computed numerically. But all these calculations need the path length which is not known. To solve this problem, this equation is used :

$$y(2l) = \frac{D}{2} \tag{25}$$

that makes it possible to compute l according to D and K_{max}.

3.2 *Timing law*

This section deals with the compactor kinematics along the path which determines the timing law of the motion. According to the compactor task, different timing law are used :

- Acceleration,
- Braking,
- Constant speed.

For each timing law, acceleration must be continuous in order to respect physical characteristics of actuators. A polynomial of the second degree is used to respect the constraints of continuity. For an acceleration from 0 to V_{max}, the velocity law is given by :

$$v(t) = -\frac{1}{\tau^3} V_{max}(2t - 3\tau)t^2 \tag{26}$$

$$\tau = \frac{3V_{max}}{2\gamma_{max}} \tag{27}$$

with :

- V_{max} is the top speed of the motion,
- γ_{max} is the top acceleration of the motion,
- τ is the duration of the acceleration.

For the braking timing law, the acceleration is inverted, but the idea is the same. Generalized coordinates of the motion can be computed by combining the path and the timing law to generate inputs of the inverse kinematic model under constraints.

3.3 *Generalized coordinates computation*

The motion is specified by the path curvature function of its arc length and by the velocity function of the time. The path curvature can be computed as a function of the time :

$$s(t) = \int_0^t v(u)\mathrm{d}u \tag{28}$$

$$K(t) = K(s(t)) \tag{29}$$

Inputs of the inverse kinematic model under constraints can be computed :

$$\omega(t) = K(t)v(t) \tag{30}$$

$$\eta(t) = \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \tag{31}$$

So the generalized coordinates are computed :

$$\begin{cases} \dot{q}(t) = S(q)\eta \\ q(t) = q_0 + \displaystyle\int_0^t \dot{q}(u)\mathrm{d}u \\ \ddot{q}(t) = \dfrac{\mathrm{d}\dot{q}}{\mathrm{d}t} \end{cases} \tag{32}$$

The compactor motion is entirely defined by joint position, velocity and acceleration. The generalized coordinates have been computed with four parameters, the width (D) and the maximum curvature (K_{max}) for the path and the top speed (V_{max}) and the top acceleration (γ_{max}) for the timing law. The Dynamic model is used to tune this parameter in order to respect the capacity of actuation and the limit due to the drum-soil interaction.

3.4 *Dynamic tuning*

It is possible to compute motor torques and drum-soil interaction forces with the inverse dynamic model. In this equation :

$$M(q)\ddot{q} + H(q, \dot{q}) = \mathrm{L}^a + \mathrm{L}^f + J^T(q)\lambda \tag{33}$$

the *Lagrange multipliers* $\lambda = [\lambda_1 \ldots \lambda_4]^T$ correspond to tangential contact force between drums and soil such as the pure rolling and non-slipping constraints are respected; the generalized actuation vector L^a correspond to motor torques. As drum-soil interaction forces and motor torques are computed, it is possible to analyze them and to tune motion parameters to optimize them.

4. SIMULATION RESULTS

Two sample motions to test this method have been used. The path is a 1.5 meter width track changing which is a typical compactor path (see section 3.1). Parameters of these motion are set in tables 2 and 3. The motion described in table 2 is a base motion, the motion described in table 3 is the base motion which has been optimized with our method.

Track width	$D = 1.5$ m
Maximum curvature	$K_{max} = 0.15$ m^{-1}
Top speed	$V_{max} = 1.4$ m.s^{-1}
Maximum acceleration	$\gamma_{max} = 3$ m.s^{-2}

Table 2. Motion planning parameters

Track width	$D = 1.5$ m
Maximum curvature	$K_{max} = 0.17\text{m}^{-1}$
Top speed	$V_{max} = 1.4$ m.s^{-1}
Maximum acceleration	$\gamma_{max} = 1.2$ m.s^{-2}

Table 3. Optimized motion planning parameters

These motion have been used as a reference in our compactor simulator. Simulation results are presented in Fig. 5 Fig. 6 and Fig. 7.

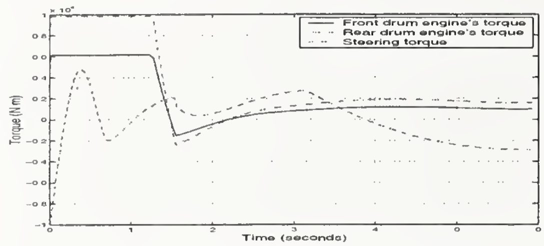


Fig. 5. Actuator torques for the base motion

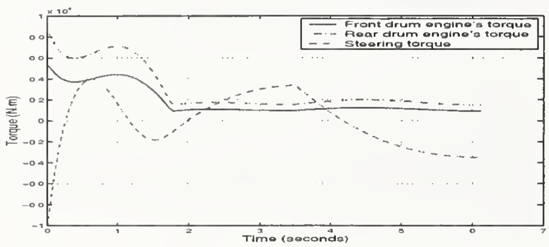


Fig. 6. Actuator torques for the optimized motion

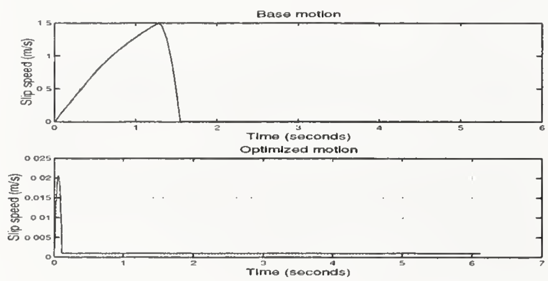


Fig. 7. Actuator torques for the optimized motion

These results show actuators torques which saturate and reach motors characteristics limits (Fig. 5). With our motion planning, it is possible to avoid saturation of actuators (Fig. 6). This optimization of the motion leads also to limited slips of the compactor drums on the soil (Fig. 7) which implies a better quality of asphalt compaction. In the frame of compaction automation, this optimized motion planning allows to design more efficient control laws. Indeed, with an optimized motion, the control of the compactor will be focus on rejecting modelling errors and external perturbations (the wind, ...). While with ordinary motion, the control law will also try to follow an unfeasible motion which implies important slips on the asphalt so a lower quality for the road.

5. CONCLUSION

A motion planning which takes into account the dynamic model of the compactor is presented in this paper. Path computing for specific tasks of compaction and the computation of timing laws have been developed to this purpose. Then the dynamic model of the compactor is used to check if the computed motion respects physical limitations of the compactor.

Motion planning is a necessary step in the implementation of a dynamic control for the compactor. It uses all works done on the compactor automation in the past years, as well for modelling as for the identification of the compactor. Now, we can consider the automation process of the compactor in order to improve compaction quality.

6. REFERENCES

Canudas de Wit, C., B. Siciliano and G. Bastin (1997). *Theory of Robot Control.* Chap. 1, pp. 4–29. Springer-Verlag. London.

Gautier, M. (1997). Dynamic identification of robots with power model. In: *IEEE International Conference on Robotics and Automation.* Albuquerque. pp. 1922–1927.

Guillo, E., M. Gautier and M. Froumentin (1999). Dynamic modelling and identification of a compactor. In: *Proceedings of International Symposium on Automation in Road Construction.* Madrid, Spain.

Khalil, W. and D. Creusot (1997). Symoro+: A system for the symbolic modelling of robots. *Robotica* **15**, 153–161.

Khalil, W. and J.-F. Kleinfinger (1986). A new geometric notation for open and closed-loop robots. In: *Proceedings of IEEE International Conference on Robotics and Automation.* San Francisco, California. pp. 1174–1180.

Façade Cleaning Robot for the Skyscraper

Thomas Bock, Alexej Bulgakow, Shigeki Ashida

The Chair for Realization and Informatics of Construction
The Faculty of Architecture, Technical University of Munich
Arcisstr. 21 80333 Munich, Germany, Tel. +49-89-289-22100, Fax. +49-89-289-22102
Email: Thomas.bock@bri.ar.tum.de : bulgakow@bri.ar.tum.de : ashida@bri.ar.tum.de

ABSTRACT: Considering all available basic conditions, a semiautomatic cleaning version can be best implemented under technical, organizational and economic constraints. All cleaning processes are executed by means of a service robot. The robot is controlled by just one operator. The cleaning quality of surface remains constantly excellent and the cleaning efficiency increases significantly.

KEYWORDS: Cleaning Robot, Façade cleaning, Semiautomatic, Vacuum suction, Nacelle traveling, Skyscraper

1. INTRODUCTION

The façades of skyscrapers have to fulfill above all a representative function for the companies. The appearance of the surface of facades plays thereby for the entire corporative identity. The environmental impact results in an intensive contamination of facades. The modern building materials react quite differently to influences of the climate. Even high-quality natural stone facades attract visible tracks of the pollution. Therefore facade surface must be cleaned regularly. The frequency of the cleaning can be executed in the cycle of decades depending on degree of pollution. The cleaning cycle and intensity are determined by economic constraints. Recently self-cleaning glass facades are available. But as far as existing façade are concerned a reduction of labor cost share by robotic maintenance is required. The higher the degree of pollution is, the more intensive cleaning agents and instruments must be used or shorter cleaning intervals.

The first skyscraper with anodized aluminum facades was built in the seventies in Munich. These facades were offered as maintenance-free. The cleaning was not considered necessary, because the anodic oxide coating is hard and weatherproof. The experience showed however that anodized facades get dirty by environmental strains. A further important argument for cleaning was not considered likewise; because the more dirty a facade is, the higher is the corrosion grade. Additionally the contamination increases the original surface, which makes it more absorptive and vulnerable for larger quantities of corrosion stress. By neglecting cleaning the façade can deteriorate as far as it will be impossible to keep its original quality. Therefore repetitive low impact maintenance by robotic cleaning anticipates real estate valueless.

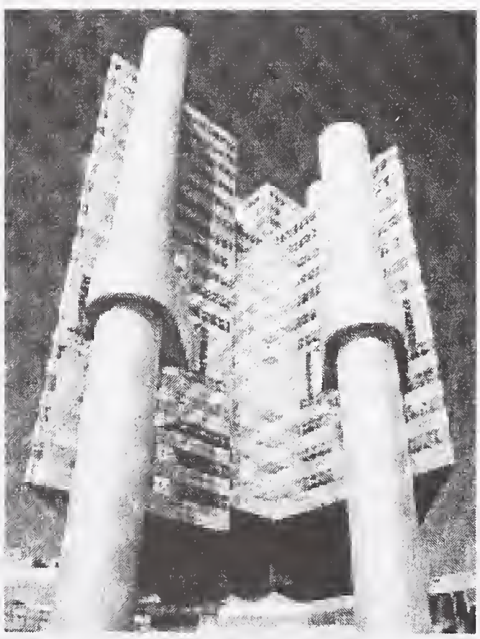


Photo 1. Skyscraper with anodized facades

Cleaning of anodized facades by robots is required for the following reasons:

- Reduction
  - of chemical pressure
  - of contamination
  - of structural changes of the original surface
- Cost-profit optimum
  - of maintenance
  - of façade appearance
  - of real estate value

2. Kinematic Solutions

2.1. The cleaning operation

The basic cleaning of an anodized aluminum surface takes place in two levels: Pre-purifying and main cleaning. Pre-purifying consists in spreading the cleaning agents on the surface and washing off with soft brushes. For the available exterior wall sections it is optimal to use rotary soft brushes. At the same time three exterior wall sections are processed. For the setbacks between the exterior wall sections brushes of an appropriate diameter must be used. These three brushes that are propelled with an electrical servo actuator form the first part of the cleaning and into those one another current waves are accommodated. At the start of the purification process three brushes are inserted into the cracks and three others received meanwhile the contact with the lamella surface. The contact pressure of rotary brushes to the surface can be determined before the cleaning and depends on the degree of pollution of the surface. The solvent is supplied by the conditioning plant and led back by a special suction procedure completely into the system. Thereby a closed cycle is implemented.

The main cleaning of the surface with a cleaning agent is executed with the help of the second section of the cleaning (Fig. 1). The brushes are exactly the same arranged, whereby for the anodized aluminum surface fine wool is used. Both brush series are arranged to each other with a distance from 30 cm. Washing the cleaning agent which arrears off must take place right after the main cleaning, so that the waste water does not dry up. In addition the third series of cleaning brushes are used. The module with two brush series can be changed, in order to execute the job of a thin water-repellent protective film with following polishing out.

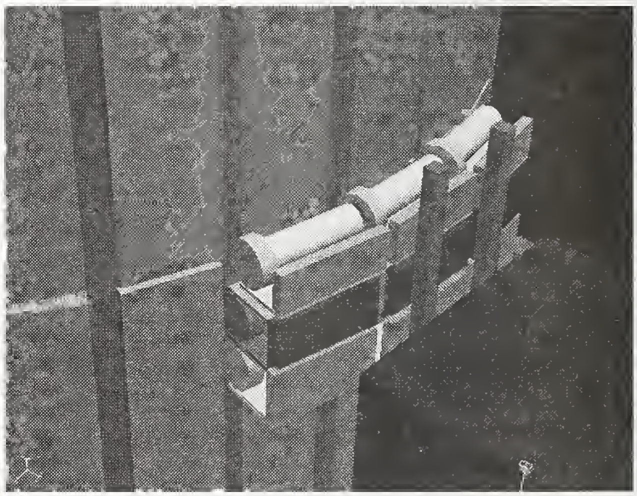


Fig. 1. Brushes in the Main Cleaning Step

2.2. The robot platform

The mobile platform of the facade-cleaning robot serves as carrier system for the cleaning operation. The necessary contact pressure to four telescope rods with vacuum plate to the facade surface is ensured at the cleaning process. Additionally they enable an occurring of the back-up ring constructions of the towers (Fig. 2).

Beside the cleaning brushes on the platform the spray nozzles are arranged for the job of a water-repellent protective film. This platform does not need guide rails at the facade, presupposed the robot autonomously, without nacelle. The progressive movement of the robot can take place via a successive separation and new positioning of each of the four suction cups. The switch-selectable bars form a basic yoke construction for the robot. This construction is completed by two rails for the progressive movement of the cleaning operation. All components of the kinematics are coordinated. The robot can be operated in automatic and semiautomatic process and used at different facades. The load entry into the facade is small. The two-coordinate construction of the robot can be served

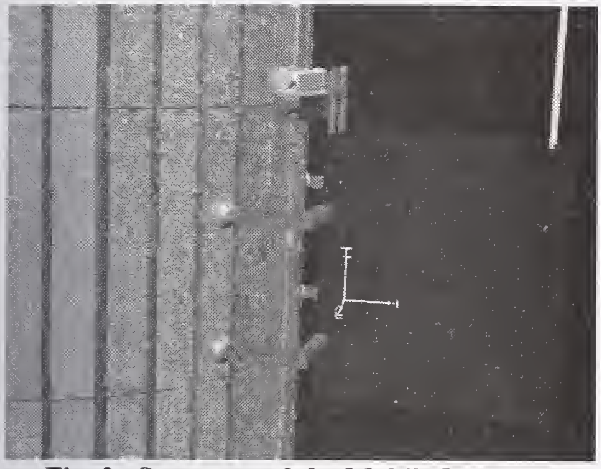


Fig. 2. Structure of the Mobile Platform

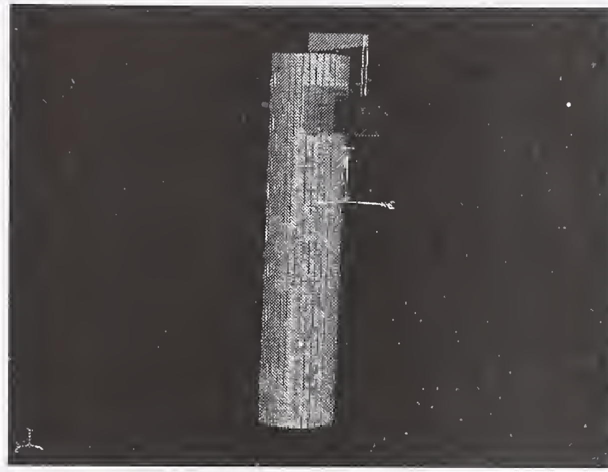


Fig. 3. Nacelle of the traveling Unit

quite easily. This kinematic pattern of the facade cleaning of robot presupposes itself a problem-free in the fulfillment of safety requirements.

3. Applying concept

3.1. Implementation

Considering all of available basic conditions, a semiautomatic cleaning version can be best implemented under technical, organizational and economic request. The robot is accommodated thereby on a nacelle of the traveling unit, adjusted while the whole cleaning process to complete and with the transfer of a user (Fig. 3). The function of the user exists in the monitoring of the cleaning process and in the execution of the cleaning of the facade surface in difficulty accessible zones by hand.

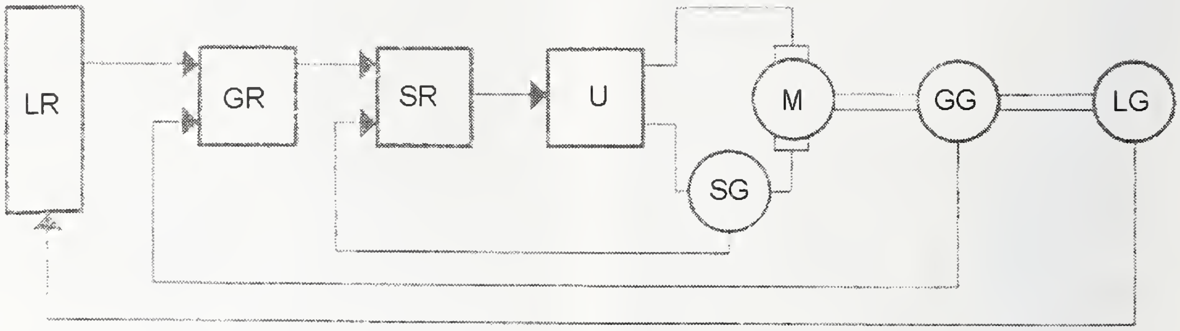
The latter enables a simplification of the robot system and thus a cost saving.
The robot can be used also at surfaces of other type, e.g. for the window surface cleaning.

3.2. Application of the cleaning operation

When an electromechanical system (power supply, engines, transmissions) is designated for the drive of the cleaning, it activates the cleaning brushes. The main parameters of the drive are performance, rate, precision, controllability, dimensions, weight and price. The special feature of the electric drive consists of the fact that they are indicated particularly by the small dimensions, an easy control, high precision, maximum stress and reliability (Fig. 4).

Table 1. Technical data of the façade cleaning robot

Type of cleaning	Brushes
Brush size	Length: 300 mm Thickness: 50 mm
Drive	Automatic guidance of the cleaning heading
Heading pressure	Manual adjusts
Dressing plant	Negative pressure system
Capacity	4,0 kW
Cleaning performance	50 m²/h
Advantages	Small personal expenditure Good and safe cleaning quality
Weight	75 kg



LR ; Position regulator, GR ; Speed regulator, SR ; Power regulator,
U ; Converter, M ; Motor, LG ; Position sensor, GG ; Speed sensor, SG ; Power sensor

Fig. 4. Main Scheme of electrical operation

Regardless of the relatively low performance and quite high price, the electric motors at the best fit for the cleaning robot of the facade.

3.3. Application of the robot platform

The platform of the robot is equipped with a pneumatic drive. The vacuum cups serve above all in order to protect the contact pressure of the robot to the surface during the cleaning process. Additionally they permit a progressive movement in up and downwards along the facade. The robot is held by the wire rope of the traveling unit. The vacuum attraction is stopped normally as a function of the entire mass of the robot, either by the degree of the vacuum or by the number of suction cups. These adjustments will be made before beginning of the work and do not need to be corrected during the job.

Adaptive vacuum cups react completely differently. At the Fig. 5 an appropriate structural variant is represented. The vacuum cup consists of the pneumatic cylinder (1), where the lower output the suction cup is in form of an open vacuum chamber (2). The vacuum chamber is connected through the connecting piece (3) with a vacuum pump. The entire device is connected by the coupling (5) with the lifting machine. The axle (4) forms a link connection to the piston rod (6) of the cylinder. The vacuum chamber has additionally a valve (7), which is a connection to the outside. Between the housing of the pneumatic cylinder and the coupling the torsion spring (8) is arranged. When releasing the instruction for touching the object, the valve is closed. In the vacuum chamber a vacuum is created by the compressor. If the load (e.g. because of the wind conditions) exceeds the nominal value, the housing of the pneumatic cylinder shifts relative to the coupling downward, the torsion spring stretches, the piston space of the cylinder becomes larger, so

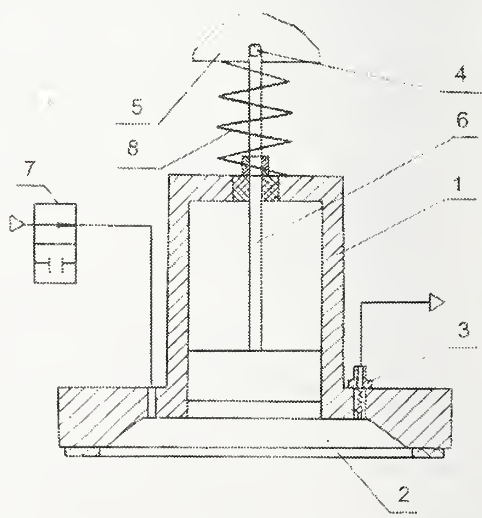


Fig. 5. Adaptable Vacuum Sucker

that the vacuum in the vacuum chamber correspond to the load of the robot increases. Suction force increases, and holding of the robot becomes more reliable. The more largely the difference between the nominal and the actual value of the load, the larger the shifts of the housing of the cylinder, the volume of the piston space and thus the suction force of the grip arm.

3.3. Control concept

The facade-cleaning robot can be differently served. In the case of a semiautomatic version the adjustment of the robot is executed in the working position and monitoring of the trial process of a user, who is with the robot in the nacelle of the traveling unit. The robot can be controlled also by means of a remote maintenance (Fig. 6). The parameters of the cleaning are controlled thereby with a monitoring camera and different sensors. The user has the function to start the purification process and react if necessary to occurring messages. A permanent monitoring of the system by the user is not necessary. However the use of a monitoring camera increases substantially the price of the cleaning robot. The electric power supply of the robot can take place by cables or via a cable attached on the nacelle.

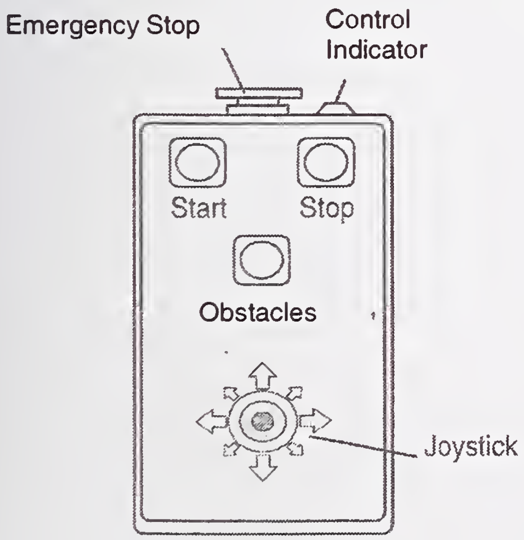


Fig. 6. Remote Controller

4. CONCLUSIONS

Considering all of available basic conditions of technical, organizational and economic constraints, semiautomatic cleaning version can be best implemented. The cleaning process is executed by the robot. The check of the process and adjusting the robot are implemented by an operator. The second worker can be saved. The quality of the surface remains constant and the cleaning efficiency increases significantly.

DEVELOPMENT AND APPLICATION OF AUTOMATED DELIVERY SYSTEM FOR FINISHING BUILDING MATERIALS

Doyama Atsuhiro, Hamada Koji, Kaneko Tomoya, Shiokawa Takashi

Building Construction Engineering Department, Technical Research Institute, OBAYASHI Corporation http://www.obayashi.co.jp/
4-640 Shimokiyoto, Kiyoseshi Tokyo 204-8558,Japan
E-mail; doyama@tri.obayashi.co.jp

Abstract: The authors have developed an automated delivery system for finishing building materials with the aim of reducing delivery costs and improving efficiency. This system is composed of two kinds of automated equipment, and managed by a delivery scheduling system via the Internet. It has been applied to the construction of a high-rise building in Osaka City, where it performed all delivery tasks with high efficiency. Furthermore, the delivery scheduling system reduced the time spent by all foremen and managers on delivery work. This paper describes functions and features of the system and some application results.

Keywords: Automation; Delivery Schedule; Finishing Work; High-rise Building; Internet

1. INTRODUCTION

Delivery of finishing materials and equipment via a construction lift or elevator (henceforth, EV) is a bottleneck in construction of high-rise buildings. It is therefore very important to improve delivery performance. Moreover, transportation does not add value to a building, so it is important to evolve its mechanization and automation.

Therefore, an automated delivery system was developed for a 56-story high-rise building in 1995 [1]. As a result of its application, the following problems remained: 1) to enable the system's application to a smaller construction site in an urban area, 2) to reduce the number of temporary EVs by using the automated delivery system even at midnight, and 3) to reduce the manpower required to input data into the system.

Delivery efficiency would be increased by solving these problems. Furthermore, by reducing the number of installed EVs, it is possible to cut down delivery costs. We have been improving the automated delivery system since 1997. During this period, partially improved systems or individual pieces of automated equipment have been applied to more construction projects.

2. GENERAL STRUCTURE OF AUTOMATED DELIVERY SYSTEM

A new automated delivery system has been developed, and was applied to the construction of a 24-story high-rise building in Osaka City in 2000, as shown in Table 1. This new system is composed of two kinds of automated equipment: an automated device that can transfer materials to an EV (henceforth, Transfer Equipment) and an Automatic Guided Folk-lift (henceforth, AGF). It is managed by a web delivery scheduling system via the Internet (henceforth, WSS). Its structure is shown in Fig.1.

A multi-stage rack suitable for a narrow space was adopted that enabled materials to be easily removed from every rack. As a result, the area of the storage facility was decreased to about 30 percent that of the previous application. WSS can access the Internet, which enables the EV to be applied without an operator needing to visit the construction site. By sharing the schedule information among several sub-contractors, it is possible to allocate the EV's hours of use. The Internet was used to obtain the input from the sub-contractors' office, also achieving labor saving.

Table1 Specifications of The Applied Building

Location	Osaka Japan
Building Use	Office Building

Structure and Size	Steel-Frame B3F/25F/PH3F Total Floor Area: 64,621□
Construction Period	Mar.1998□Oct.2001

3. AUTOMATED EQUIPMENT

The authors planned to cut down the number of installed EVs by applying the automated delivery system. We estimated the overtime work at a maximum of 4 hours per a day during the height of finishing work. The automated equipment was mainly used for delivery during overtime work for basic materials such as gypsum plasterboard and lightweight steel.

The layout of the ground floor used as a storage facility is shown in Fig.2. Trucks carrying materials enter the building and the materials are unloaded onto the floor. The AGF operates automatically from its home position, and then unloads the materials onto the multi-stage rack. During the automated delivery work, the AGF automatically removes the materials from the multi-stage rack and conveys them to the front of the EV (HCE2800 in the figure). AGV's operations are repeated according to the operator's pre-inputted orders. After the automatic Transfer Equipment in the EV loads the materials, the EV is taken up to the delivery floor by an elevator operator. When it arrives at the delivery floor, the automatic Transfer Equipment automatically unloads the material from the EV.

3.1 AGF and Multi-stage Racks

The AGF's rated capacity is 1500 kgf. It runs along an electromagnetic guide wire. This guidance system is more reliable than others such as light reflex or magnetic tape. An AGF removing material from a multi-stage rack is shown in Fig.3. We reduced development costs by adapting a standard machine to an AGF by attaching a sensor and radio equipment needed for construction work. If the AGF detects any obstacles in the traveling direction with its photosensors or supersonic sensors, it begins to slow down to avoid contact with them. Moreover, the AGF and Transfer Equipment have a collision avoidance system that utilizes radio signal communication. The AGF can deliver materials safely and efficiently in a narrow space. We applied 2m- and 4m-wide

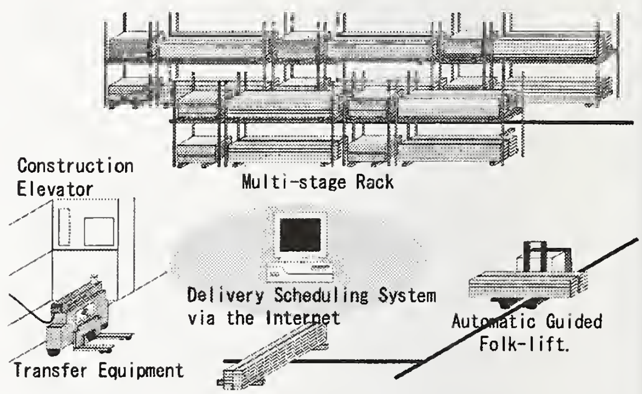
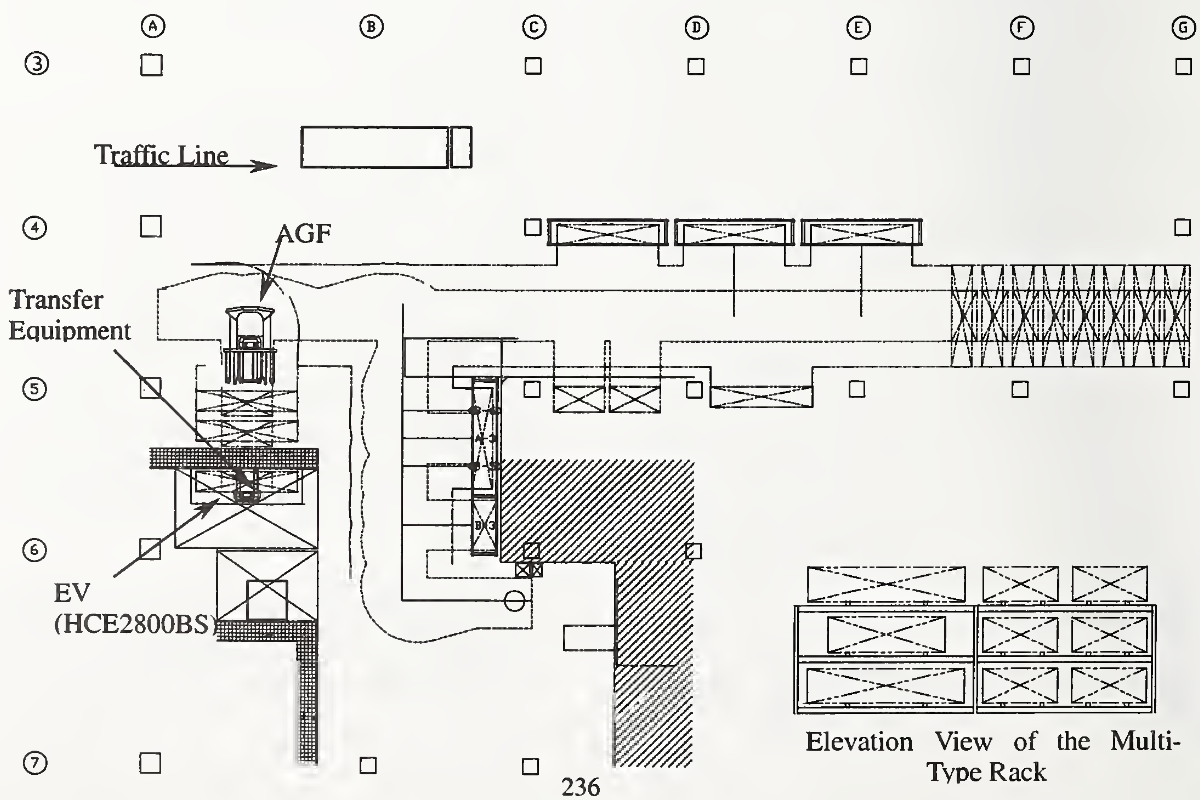


Fig.1 The Structure of Automated Delivery System



racks adapted to the form of object materials. An address number was set for each rack. After the operator inputs the address number into the AGF's control panel, it continuously delivers materials up to 99 times.

3.2 Automatic Transfer Equipment

The Transfer Equipment shown in Fig.4 is attached to an EV. Materials are loaded and unloaded automatically by an elevator operator. By automating the unloading work, it is possible to reduce delivery time and achieve labor saving. The Transfer Equipment is as light as 800 kgf, whereas its rated capacity is 2000 kgf. It is supplied power directly from the EV via a flexible cord, so it does not need to carry a heavy battery. Moreover, when an EV is used to transport workers or lightweight materials, the Transfer Equipment can be easily removed from it in order to maximize its capacity.

4. WEB DELIVERY SCHEDULING SYSTEM

4.1 Structure of Web delivery scheduling system

The structure of the Web delivery scheduling system (WSS) is shown in a Fig.5. It is composed of three kinds of hardware: a TA-terminal for application; a WS-Web shared server; a TM-terminal for management. The TA is a computer installed in each subcontractor's office, which accesses the Internet and processes several assignments: new reservations, deletions, modifications, and confirmations for use of an EV. The WS implements script in response to requests from the TA. Since a CGI pearl is used for script, a hosting service provided by a general Internet service provider can be used. The TM is a computer installed in a construction office,

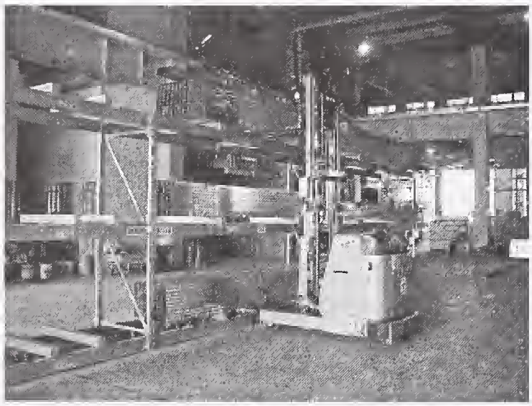


Fig.3 Automatic Guided Folk-lift

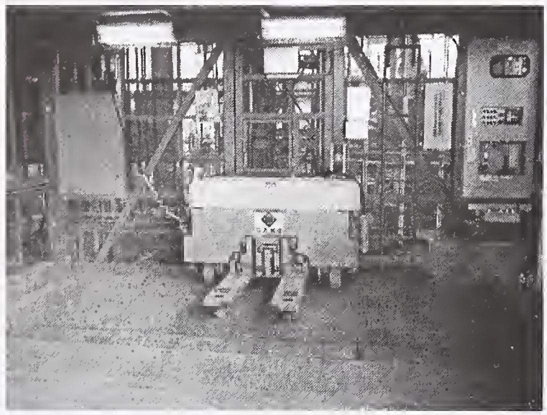


Fig.4 Transfer Equipment

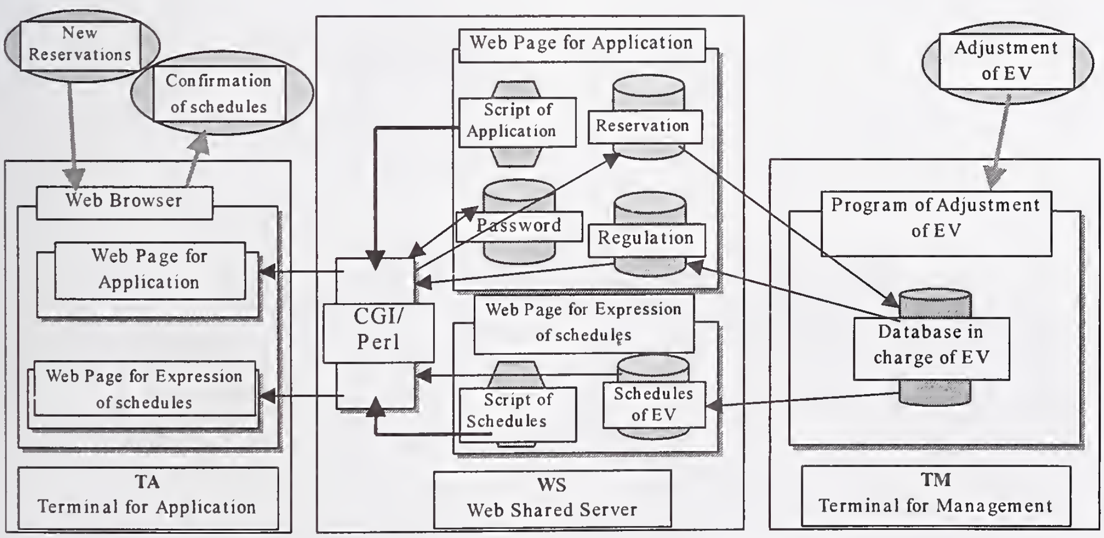


Fig.5 The structure of The Web Delivery Control System

which downloads data for a definite period and makes a schedule based on a predetermined algorithm and uploads these data to the WS as fixed information. Moreover, the TM is operated by the EV's manager and makes EV regulations such as holidays and regular services, and uploads these data to the WS. Consequently, an EV reservation can be restricted. In addition, the TC can indicate past records of an EV in contrast to plans. Fig.6 is a calendar indication that shows application conditions. Detailed application conditions obtained from the bar chart can be confirmed by choosing the reservation day (Fig.7).

4.2 Management of Transfer Web delivery scheduling system

The EV's manager and subcontractors held a meeting once a week to adjust the EV schedule for the following week. Before the meeting, the EV's manager downloaded the application data and closed the applications. After the EV schedule's adjustment, the administrator uploaded the determined schedule to the WS.

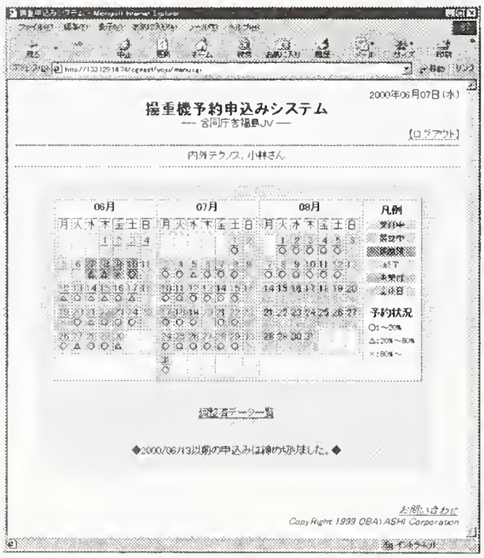


Fig.6 Calendar Indication of WSS

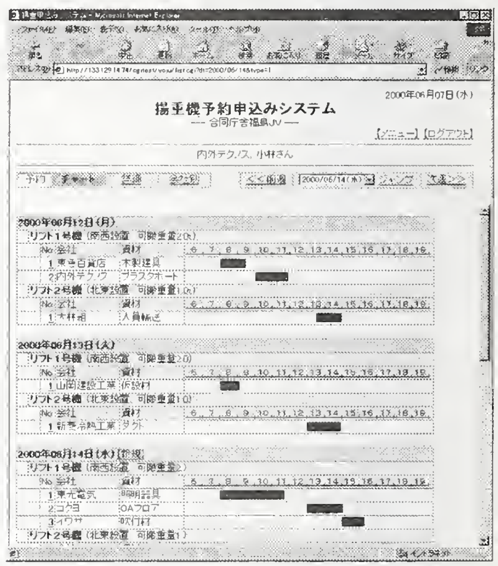


Fig.7 Bar Chart Indication of WSS

5. APPLICATION RESULT

5.1 Automated Equipment

We analyzed the EV's operating rate - Or - described by:

$$Tn = To - Ti$$
$$Or = Tn / To \qquad (1)$$

Where: Tn is net EV operating time, To is EV operating time, and Ti is EV downtime (waiting for materials, workers, etc.).

Fig.8 shows the EV's operating rate transition every day during the operating period. Fig.8 contrasts the Or of this system with that of the conventional system for the same construction scale. Fig.8 indicates that

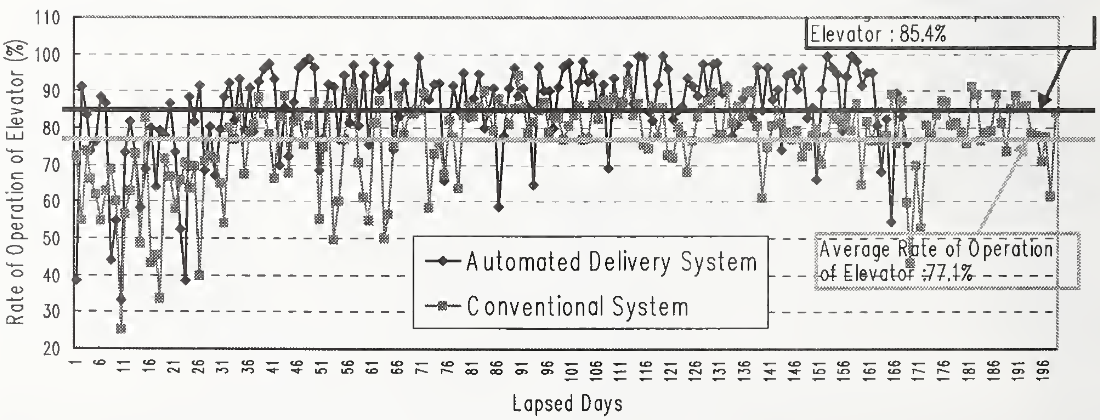


Fig.8 The Transition of EV's Operating Rate

the automated system average was higher than that of the conventional system.

Fig.9 compares the delivery performances of the conventional and automated systems. The conventional system's performance is assumed to be 100. The average weight between the two systems is approximately equivalent, whereas other items of the automated system exceed those of the conventional system. In particular, the automated system shows a delivery manpower saving of nearly 45 %. By comparing the two systems, we confirmed that the delivery capacity per unit time of the automated system was 1.44 times that of the conventional system.

It thus follows that the automated delivery system is more efficient than the conventional system. The main reasons are:

1) The application of full-time material transportation workers - these workers can deliver materials stocked on multi-stage racks in the EV's spare time
2) Setting of multi-stage racks - this enabled the EV to deliver materials even when there was a delay of delivery trucks
3) Introduction of two kinds of automated equipment - these enabled a reduction of delivery work and labor saving

5.2 Web delivery scheduling system

We analyzed logs stored in the server of the Web delivery scheduling system for seven months from August 2000 to March 2001. Fig.10 shows the transition of the number of log-ins according to the type of business. This shows that the total number of log-ins was 2700, and the average weekly logins was 84. Furthermore, the number of log-ins decreased gradually with time. The number of log-ins according to the type of business confirms that the percentage for building equipment was 55.4 %, and the percentage for building finishing work was 26.4 %.

Fig.11 shows the percentage of access points to the WSS. As shown, the access points to the WSS were mostly out of the construction office.

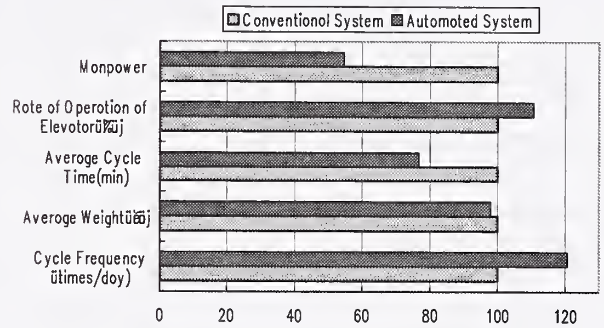


Fig.9 The Comparison of Delivery Performance

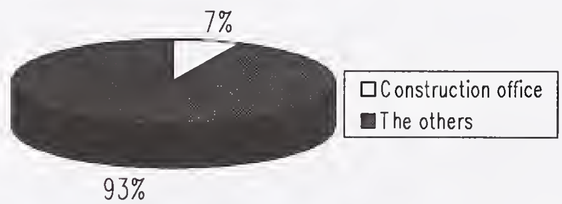


Fig.11 The percentage of access points to the WSS

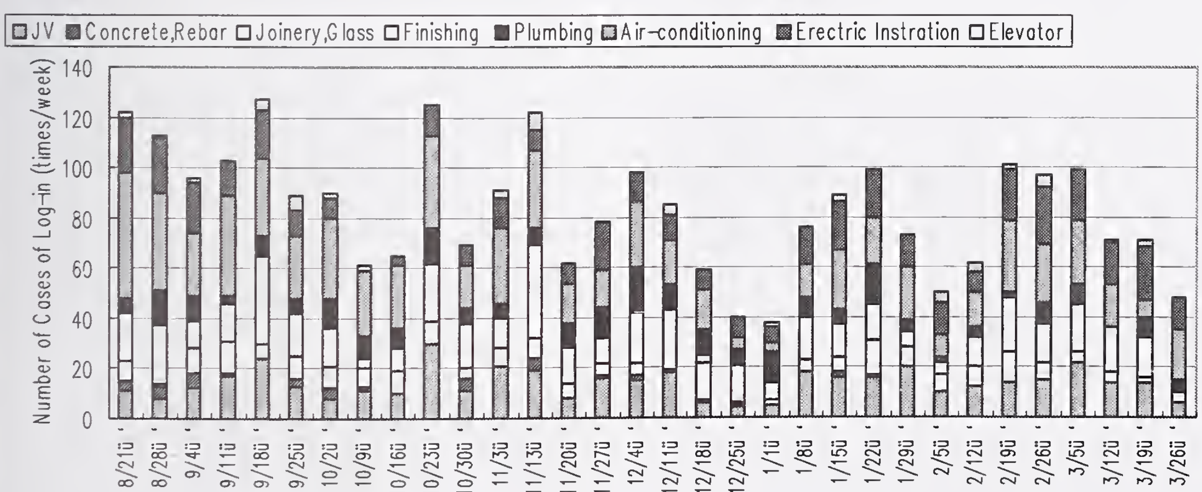


Fig.10 The Transition of The Number of Log-ins

Fig.12 shows the number of log-ins according to the time period. As shown, reservations during regular working hours (8 a.m. -5 p.m.) occupied 60 % of the total, whereas those during holidays or from midnight to early morning occupied 40 % of the total.

We carried out a questionnaire survey of about twenty users of the Web delivery scheduling system. The findings of this questionnaire survey include:

1) Although there were many inexperienced users, the average log-in time was short, about ten minutes, and the operability and speed were highly rated.

2) There were many positive comments on the laborsaving ability.

3) The assignment, such as data input and adjustment, of the EV's manager was reduced by 20%.

In view of the results so far achieved, we concluded that it is very effective to use the Internet for EV management, because the EV can be reserved any time and from any place.

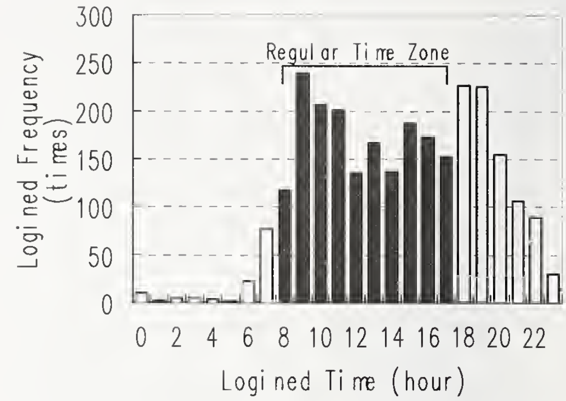


Fig.12 The Number of log-ins According to The Time Period

6. CONCLUSION

An automated delivery system has been developed for reducing delivery costs and improving delivery efficiency. Results show that the delivery capacity per unit time of the automated system is 1.44 times that of the conventional system, and it achieves a 45% delivery labor saving. In addition, using the Web delivery scheduling system reduced the manager's assignment of the EV. We covered development costs of automated equipment for one project. Results show that total delivery costs can be as little as initial costs because of the reduced number of installed EVs.

We believe that delivery costs can be reduced even further by expanding the application range of the automated delivery system. Moreover it is hoped that the automated delivery system will develop into construction logistics by improving the management method in a storage facility and expanding the WSS's functions.

References

[1] Hamada, K et al. (1996) "Development of an Automated Material Transportation System for a High-rise Building", Report of Obayashi Corporation Technical Research Institute, No.52, 113-120.

[2] Hamada, K et al. (2002) "Development of Automated Delivery System for Finishing Building Materials", Report of Obayashi Corporation Technical Research Institute, No.64, 1-6.

[3] Doyama, A et al. (2002) "Development and Application of Automated Delivery System for Finishing Building Materials", 14th Construction Robot Symposium, 19-26.

UNMANNED CONSTRUCTION SYSTEM: PRESENT STATUS AND CHALLENGES

Yoshio Ban[1]

ABSTRACT: This report analyzes past unmanned construction system done at the Unzen-fugendake and Usuzan volcanoes to organize the present status of unmanned construction systtem from varied perspectives, describes challenges and countermeasures, and introduces future prospects.
Unmanned construction systtem have been done to perform emergency countermeasure work and restoration work at disaster sites. Remaining challenges include development of methods of performing remote hillside reforestation in devastated areas, remote traveling crane operation. Turning to the future, it is proposed that noise and vibration be provided around the operator's seat in remote control rooms to create a sense of realism. Unmanned construction systtem should be used not only at disaster restoration sites, but also used effectively to increase safety at ordinary construction sites.

KEYWORDS: Unmanned construction, remote operation, hillside reforestation, traveling crane, rough terrain crane

1. INTRODUCTION

Unmanned construction is work performed by remotely operated construction machinery that corresponds to an operator controlled robot. Unmanned construction was used in civil engineering work for the first time in Japan in 1969 when an underwater bulldozer was used to excavate and move deposited soil during emergency restoration work at the Toyama Bridge that had been blocked by the Joganji River disaster. Since then, unmanned construction by excavators inside pneumatic caissons and by backhoes have been carried out, but the restoration work following the volcanic eruptions that began in 1994 at the Unzen-fugendake Volcano and restoration work executed following the eruption of the Usuzan Volcano in 2000 were the first executions of large-scale unmanned construction and have spurred rapid progress in unmanned construction technologies and encouraged their wide use.

This report analyzes recent applications of the unmanned construction method including those at Unzen-fugendake and Usuzan Volcanoes, outlines the present state of unmanned construction from multiple perspectives, discusses future challenges and measures to overcome them, and concludes with future prospects for the method.

2. ANALYSIS OF THE PRESENT STATE OF UNMANNED CONSTRUCTION

The following are analyses of unmanned construction executed at the Unzen-fugendake and Usuzan Volcanoes.

2.1 Work Categories

Past executions can be broadly categorized as emergency works executed at the time of a disaster and later restoration works. The method has often been used after disasters including those caused by debris flows and pyroclastic flows, the collapse of soil caused by earthquakes and so on.

The following are the principal categories of work performed by this method.

[1]Advanced Construction Tecknology Center, 2-15-6 Otsuka Bunkyo-ku, Tokyo, Japan (T) 03-3942-3992
(F) 03-3942-0424 E-mail ban@actec.or.jp

Table 1. Principal Categories of Work Performed by this Method

Work Category	Description of the works
Rock removal work	Excavation, loading, transporting
Structure demolition and removal work	Crushing and pulverizing concrete and cutting steel reinforcing bars, loading and transporting the products
Large sandbag placing work	Transporting and placing
Concrete block work	Removing obstructions, leveling ground, placing
Temporary road work	Cutting, filling, and compaction
Erosion and sediment control dam work	Excavation, embanking, backfilling, compaction, pouring concrete
Watercourse work	Excavation, pouring concrete, placing foot protection blocks
Tree felling work	Cutting, stumping, transporting
RCC work	Transporting, spreading and leveling, compaction, spraying, laitance removal
Ready-mix concrete work	Installing form materials, pouring and compacting concrete
Soil form work	Excavation, loading, transporting, removing form materials

2.2 Remotely Operated Construction Machinery

The following are types of remotely controlled construction machinery that have been used frequently.

Table 2. Frequently Used Remotely Controlled Construction Machinery

Machine Type	Standards
Back hoes	0.6 to 4.3 m^3
Bulldozers	16 to 80 t
Dump trucks	32 to 80 t
Crawler type trucks (large tracked trucks)	6 to 12 t

In addition, towed transport vehicles, trucks, tractor shovels, crawler cranes, self-propelled cranes, crane trucks, vibrating rollers, etc. have been used.

2.3 Unmanned Construction Equipment

The following are the types of unmanned construction equipment that have been used according to the radio transmission distance.

Table 3. State of Use of Unmanned Construction Equipment

Radio transmission distance	Remote operation room	Radio relay truck	Mobile camera vehicle	Stationary camera
Shorter than 50 m	□	□	Used according to circumstances	
50 m to 150 m	□	□		
Longer than 150 m	□	□		

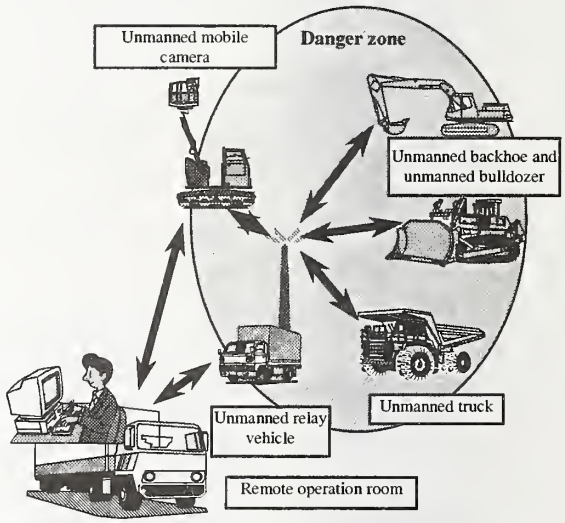


Figure 1. Configuration of Unmanned Construction Equipment

2.4 Types of Radio Equipment

The frequencies of radios used for unmanned construction are 400 MHz, 2.4 GHz, 50 GHz etc. Various innovations include lowering their transmission power to limit their range, thereby avoiding the need for licensing. Each radio base station is selected according to the transmission distance, rectilinearity, channel capacity, data transmission capacity, licensing requirement, and other differences, and according to the work environment.

2.5 Transporting, Assembling, and Modifying the Machinery

The number of days required to obtain the machinery is a problem that must be resolved, particularly when performing emergency restoration at the time of a disaster. The number of days required to acquire the machinery varies sharply between cases where machinery already equipped with remote control systems is transported to the site and cases where the hydraulic circuits of construction machinery are modified. In the past, it has taken a few days in the former cases and a few months in the latter cases.

2.6 Work Efficiency

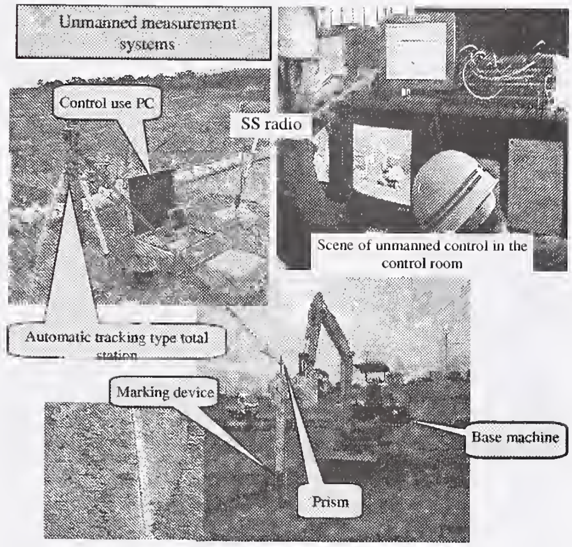
It is said to be roughly 60% to 70% of that of manned construction, but the efficiency falls sharply in cases where the machinery moves, where materials are transported in dump trucks for example, and in cases where high precision work is necessary.

2.7 Controlling and Inspecting the Construction

Work requiring extensive finished work control is now done by manned methods. The following systems are used for unmanned construction.

1) Automatic tracking type total station
   A prism that functions as a target and a marking spray device are installed on the front of a remotely controlled back hoe, and the automatic tracking total station measures the target to mark the position to install the back hoe and to confirm its position is correct.

2) GPS precise real time positioning system
   This technology accurately computes the position of each working machine by combining information about the position of the machines obtained by RTK-GPS with status information provided by sensors installed on various parts of the construction machinery. For example, in some cases, GPS receivers and prisms have been installed on bulldozers or vibrating rollers etc. so the system can perform coordinate control by obtaining measured data to control the spread thickness or the compaction frequency.



Photograph 1. Unmanned Measurement Systems

3. FUTURE CHALLENGES FACING UNMANNED CONSTRUCTION AND MEASURES TO OVERCOME THEM

3.1 Work Categories (introduction of unmanned hillside reforestation)

It is expected that in addition to various work categories in sediment and erosion control work such as permeable dam construction or precast block masonry work, unmanned executions will play an important role in hillside work in devastated areas carried out to prevent disasters and to conserve the environment.

Hillside work is broadly categorized as hillside foundation work and hillside reforestation work. Hillside foundation work is done to lay the foundation for a future forest by stabilizing deposited sediment, while hillside reforestation refers to introducing trees directly on hillsides to form forests that will prevent the runoff of sediment. Execution methods that can be used for the former, hillside foundation work, are concrete block masonry work, placing large sandbags, and placing boulders at the site: methods that can be executed with existing technologies. The latter, hillside reforestation, has been done as unmanned reforestation spreading work: spreading seeds and other materials with a high speed belt conveyor type spreading machine. And although it has not been done unmanned, seeds have been sown from the air by manned helicopters for several decades.

Photograph 2. High Speed Belt Conveyor Type Spreading Machine

But on devastated ground where the soil condition is poor and the sediment moves, reforestation by transplanting seedlings is more reliable than seeding work, so there is a demand for unmanned transplanting methods.

- Hillside work
  - Hillside foundation work
  - Hillside reforestation work
    - Seeding
    - Transplantii

The wood frame block root pot method has been proposed as an unmanned transplanting work method. Seedlings are grown inside wood framed blocks consisting of a wooden frame, (approx. 1.0 m•1.0 m), net (palm fiber etc.), water holding material, and leaf mold, and the frames are placed in rows by an unmanned backhoe equipped with a holding device. This method does not require separate works such as digging holes and mulching, and it can be used on any slope where a backhoe can be operated. It should be introduced as a working system in the future.

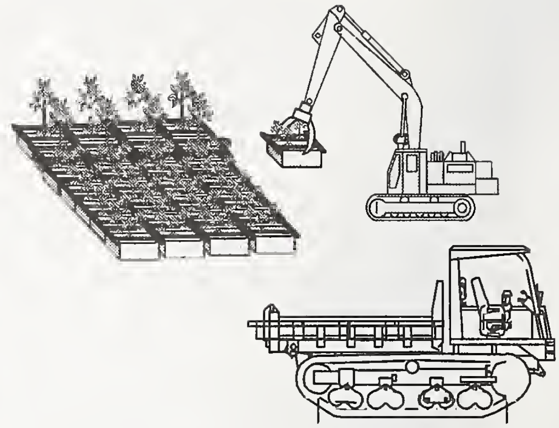


Figure 2. Image of Unmanned Transplanting

3.2 Remotely Operated Construction Machinery (introduction of unmanned traveling cranes)

In June 2002, the Ministry of Land, Infrastructure and Transport of Japan announced that it had begun to develop unmanned operating technology for traveling cranes. Till now, medium and large traveling cranes have never been remotely operated and developing the world's first successful unmanned operating system for large cranes will be a great achievement. Because cranes may be installed in narrow places and on slopes, the crane chosen for the development project was a rough terrain crane, a type that can be kept in a horizontal attitude with an outrigger and handled in narrow spaces. (lifting capacity: 50 tons)

A rough terrain crane, which is a type of wheeled crane that is propelled by a single engine and can perform crane work, travels easily over uneven ground because it is a four wheel drive vehicle and all four wheels can be steered. Traveling and crane work are controlled from the same operating cab. Although externally it appears more compact than a track crane, it provides comparable lifting capacity. In Japan, therefore, most cranes manufactured are traveling cranes.

The first step in the development is to decide which traveling crane functions will be performed remotely. Because cranes work while installed in dangerous locations, almost all traveling and crane operation functions must be remotely operated. Because unmanned construction is performed rarely, it must also be possible for the crane to be operated by an operator sitting in the cab when it is used in safe places.

The next problems is attaching and detaching loads. It is extremely difficult to attach loads to and detach them from the hook by the remote control method. An automatic load detachment device operated by a weak radio wave has already been developed and is in use on small cranes. The following two solutions are being considered. One is to replace hooks with a newly developed device that attaches a load by grasping it. The second is to limit attachment work to manned areas. (If attachment work is limited to manned areas, the working range of a 50 t class rough terrain crane will be about 60 m from the boundary of the manned and unmanned areas.)

The next problem is the "prohibition on leaving the operating location." Japanese law includes a provision that states, "The operator of a traveling crane shall not leave the operating location while a load is still suspended by the crane" (Regulation for Safe Operation of Cranes Etc.). Therefore, to comply with the Regulation for Safe Operation of Cranes Etc., an operator's seat must be provided in the remote control room and it must be equipped with a device that can provide the same information available at the operator's seat in the crane. Specifically, it is necessary to provide a system that converts the boom angle, rotation angle, and other quantities representing the state of the crane, overload warnings, and visual information about the load and the wire (visual data) into electronic data and transmits this data to the operating room.

When unmanned crane work has been realized, it is counted on to be applied to various kinds of work in dangerous locations. Expanding the working radius of unmanned construction will not only permit execution of construction work in previously inaccessible locations; it is also counted on to permit the installation of larger blocks, increasing work efficiency.

4. FUTURE PROSPECTS

4.1 An operator's Seat with a Sense of Reality

When unmanned execution is too far away for the operator to see the site, the operator remotely controls the execution in a remote control room while watching images transmitted from television cameras. Even when relatively simple work such as backhoe or bulldozer work is executed, the operator uses two or more television cameras that provide views of the overall scene and the view from the cab of the machine. To perform work requiring detail operations or depth perception, special measures are taken at each site, installing another camera providing a view from the side.

But during manned work (when the operator sits in the seat in the cab of the construction machine), the operator operates the machine while unconsciously obtaining information other than visual information (noise, vibration, deviation from the horizontal, etc.), and this difference is the cause of gaps in execution efficiency, execution precision, and reliability between manned and unmanned work.

In the future, the operator's seat must be made more realistic by adding noise, vibration, and other information useful for operators so that they can make the maximum use of their five senses.

4.2 Changeover to Safer Construction Methods

Unmanned executions can be divided into the following three stages.

(1) Emergency measures work at a disaster location
(2) Restoration work at a disaster location
(3) Ordinary work at dangerous locations

Now that unmanned construction has been firmly established and has gained a good reputation as disaster site restoration work method, it is expected to advance to the next step: as a method of achieving safer work at normal construction sites.

The following are considered to be dangerous places to execute civil engineering works.

- Slopes at risk of failure (sediment control dam work, hillside work, slope work)
- High places (slope work, concrete dam work)
- Riverbeds and lake bottoms (dredging work)
- Seabed (dredging work, leveling riprap)
- High pressure environments (pneumatic caissons)
- Places at risk of a gas explosion (vertical shaft excavation)
- Places exposed to radiation or dioxin and other hazardous substances (structure demolition and removal work)

Adopting unmanned construction for work at such locations can be counted on to have the following effects.

- Improved safety
- Reduction of temporary work costs and safety measure costs
- Improved operating rate (shortening construction periods)
- More precise higher quality construction

For example, seabed dredging has been done from the surface using a conventional crab barge or pumping barge, but if dredging is done by an unmanned machine traveling on the seabed, the work can be done safely during heavy seas, sharply increasing the operating rate (work period) and execution precision (quality). This technology has already been used for more than ten years remotely controlled by a wired system.

5. CONCLUSION

The development and popularization of unmanned construction will have many effects: improving the safety of construction work at the same time as it shortens construction periods and guarantees high quality. This will, in turn, improve the image of the construction industry that tends to be negative in Japan and spur progress in IT and robotics.

In the future, we wish to accurately assess the present state and challenges facing unmanned execution technologies and work to popularize their use.

REFERENCES

1) Advanced Construction Technology Center (ACTEC) (editor): Handbook on the use of unmanned construction in emergencies (published in August 2001)

2) Ito H.: Understanding traveling cranes (published in December 1994)

Report of the NIST Workshop on Automated Steel Construction [1]

by

Alan M. Lytle[2], Kamel S. Saidi[2], William C. Stone[2], and John L. Gross[2]

ABSTRACT: The Building and Fire Research Laboratory of the National Institute of Standards and Technology, in cooperation with the American Institute of Steel Construction, sponsored a workshop on automated steel construction. The purpose of the workshop was to investigate the development of new technologies to facilitate automating the steel construction process. Desired outcomes included a clear definition of issues and constraints, the identification of candidate breakthrough technologies, and the development of a research roadmap. A description of the workshop structure, agenda, and preliminary results are presented.

KEYWORDS: construction automation, automated steel erection

1.0 INTRODUCTION

Productivity, reliability, and safety are the three predominant issues facing the steel construction industry today. In both industrial facilities and commercial buildings, hot-rolled steel members are typically joined together either by welding or using high strength bolts. These processes require a significant amount of skilled labor, and in the case of high-rise construction, constitute one of the most dangerous specialties in the already hazardous construction industry. Inspection is difficult and time consuming, and often, the connections are the weakest link in the resulting structure.

The steel construction industry faces significant challenges to remain competitive. The following two statements succinctly summarize the issue [1]:

> "The U.S. construction industry must begin to move away from a nearly exclusive labor-intensive business and towards automation to be competitive in the ever-shrinking global marketplace."

> "Decreasing fabrication and erection time for steel frame buildings, while increasing the safety of workmen during construction are issues that must be addressed, and provides the motivation for automated construction."

According to the American Institute of Steel Construction (AISC), a 25 % reduction in time required to erect a steel frame structure is needed. In response to this stated need, the NIST Building and Fire Research Laboratory (BFRL) and AISC co-sponsored a workshop on Automated Steel Construction at the NIST campus in Gaithersburg, MD on June 6 and 7, 2002. The workshop brought together steel

[1] Official contribution of the National Institute of Standards and Technology (NIST); not subject to copyright in the United States.

[2] NIST, BFRL, Materials and Construction Research Division, Mail Stop 8611, Gaithersburg, MD 20899-8611; alan.lytle@nist.gov, kamel.saidi@nist.gov, william.stone@nist.gov, john.gross@nist.gov.

producers, fabricators, designers, erectors, and construction automation experts to discuss factors affecting the steel construction industry and to identify possible courses of action to assist the industry. The desired outcome from the workshop was a clear definition of issues and constraints, identification of candidate breakthrough technologies, and the development of a research roadmap. This report presents information contained in the keynote addresses and results of the working group breakout sessions.

2.0 WORKSHOP FORMAT

The workshop convened over a period of 2 days, during which the participants discussed the challenges faced by the US steel construction industry and the various technologies that could be used or developed in order to meet those challenges.

There were a total of 17 non-NIST attendees at the workshop. Participants included representatives from 3 steel producers, 3 steel fabricators, 2 steel designers, and 6 steel erectors, as well as 3 construction robotics and automation researchers. NIST researchers included personnel from BFRL's Materials and Construction Research Division and the Manufacturing Engineering Laboratory's Intelligent Systems Division.

The workshop was divided into three sessions over a day and a half. Each session included a topical presentation[3], a breakout session, and a full group discussion.

3.0 DETAILED AGENDA

3.1 Day One

Day one began with an introduction by the NIST Construction Metrology and Automation group (CMAG) leader, Dr. William C. Stone.

[3]All workshop presentations available on the CMAG website: www.bfrl.nist.gov/861/CMAG/index.html

The introduction was followed by a presentation by Dr. Carl T. Haas of The University of Texas at Austin entitled "Automation in Steel Erection" [2]. Dr. Haas' presentation included a definition of industry problems, possible opportunities for automation, and a review of some previous construction automation research and development activities. Specific opportunities for automation discussed included:

- Robotics and process integration in the fabrication shop
- Materials tracking using radio frequency identification (RFID) tags, bar codes, etc.
- Design of connections for compliant assembly
- Pre-assembly to minimize field connections
- Integrated project processes, databases, and 4-D models
- Positive control of members and subassemblies using manipulator arms, inverse Stewart platforms, etc.
- Automated welding, bolting, adhesion, etc.
- Global positioning and locating systems

Examples of previous applicable research and development presented included:

- Lehigh ATLSS connection [3]
- NIST RoboCrane [4]
- Japanese automated building systems [5,6,7]
- UT Large Scale Manipulator [8]

Following the first presentation, the workshop participants were divided into 4 groups. Each group was tasked with forming a list of technologies that could benefit the steel construction industry and a corresponding list of criteria that could be used to rank those technologies.

The lists of technologies from the 4 breakout groups were then presented to all the workshop participants and discussed. A single list of the most promising technologies and a single list of evaluation criteria for those technologies was then developed by the workshop participants. These lists are presented in Tables 1 and 2.

The afternoon session of day one began with a presentation by the president of National Riggers and Erectors, Inc. (Plymouth, MI), Mr. Robert Dunn, entitled "Steel Erection and Challenges" [9]. Mr. Dunn reviewed the challenges facing the steel industry including safety, quality, workforce aging, and the cost of construction. He then reviewed various elements of the erection process, presented a cost breakdown of those elements, and projected potential cost benefits of various process improvements. Mr. Dunn stated the areas with the greatest opportunity for potential cost savings include:

- Ground Operations (receiving, etc.)
- Hoisting
- Ground Assembly
- Temporary Bracing
- Detailing

In his conclusion, Mr. Dunn outlined the following 4 recommendations for application of automation to steel erection and the corresponding potential cost savings:

1. Pre-assembly and/or modularization of roof/floor/wall components can save 10 % to 20 % of ground operations/hoisting costs which constitutes 45 % of total erection cost – a 4.5 % to 9.0 % overall savings.
2. Optimizing crew sizes and using innovative lifting/storage devices can save 15 % to 20 % of hoisting cost which comprises 30 % of the total erection cost - a 4.5 % to 6.0 % overall savings.
3. Use of "snug-tight" bolts in bearing connections can realize savings of from 30 % to 35 % of this cost driver which accounts for 30 % of erection costs - a 9.0 % to 10.5 % overall savings.
4. Semi-automated welding practices can save 2 % to 5 % of overall erection cost.

Following Mr. Dunn's presentation, the workshop participants were again divided into 4 groups. Each group was tasked with forming a list of challenges that the US steel construction industry faces and a corresponding list of criteria that could be used to rank those challenges.

The lists of challenges from the 4 breakout groups were then presented to all the workshop participants and discussed. A single list of the most important challenges and a single list of evaluation criteria for those challenges was then developed by the workshop participants. These lists are presented in Tables 3 and 4.

Once the lists of technologies, challenges, and corresponding evaluation criteria were developed, each workshop participant was then asked to choose what they felt were the 5 most important technologies and the 5 most important challenges. The participants were then asked to rank their chosen technologies and challenges in terms of their relative importance to one another.

The participants were also asked to repeat the same process for the lists of criteria corresponding to the technologies and challenges; however, in this case the participants were asked to choose only three criteria from each list.

3.2 Day Two

Based on the ranking of the technologies and challenges (and their corresponding criteria from day one) day two of the workshop began with the participants scoring the top 5 technologies and challenges (and the top 3 criteria for each) using the Analytical Hierarchy Process (AHP), which is described in section 4.0.

The scoring process was followed by a presentation by Dr. Jim Ricles of Lehigh University (Bethlehem, PA) entitled "Next Generation Steel Structures." Dr. Ricles reviewed steel framing, structural requirements for connections, and current connection schemes. To establish the need for automated steel construction research, Dr. Ricles provided the following summary statements:

- Construction industry comprises approximately 8 % of the U.S. Gross National Product.

- U.S. construction productivity has shown an average annual net decrease of nearly 1.7 % since 1969.
- Procedures for erecting building structures have changed very little over the past 80 years (although rivets have been replaced by bolting and welding). Erectors perform strenuous tasks in a highly dangerous environment.
- Incidents of occupational injury reported for construction workers comprised over 10% of all cases.
- Workers' Compensation Insurance for steel workers is 19.3% of wages, the highest of all construction workers.
- Percentage of fatalities in the construction industry (limited to building erection) from falling is 43%.

Dr. Ricles then discussed required connection characteristics for automated construction and provided examples of connection ideas for automated construction. Characteristic features required of next-generation beam to column included [10]:

- *Self-alignment* – The connection must be able to guide the beam toward the proper location once contact is made between connection elements located on the beam and column.
- *Tolerances* – The connection must have tolerances which allow for misalignment or out-of-plumbness.
- *Adjustment* – Because of the tolerances that must be built in, it is unlikely that the connection will be precisely in its correct position after erection. Therefore, the connection must have the ability to be adjusted easily.
- *Stiffness, Strength and Stability* – The connection must be strong enough to carry erection loads while possessing a suitable amount of stiffness to control deflections. Furthermore, the connection must be stable enough to allow erection of the structure to continue until the final fastening.
- *Modularity* – The connection should be able to be mass-produced with a standard shop fitting operation and with quick, automatic erection capabilities.

Following the presentation, the workshop participants were once again divided into 4 groups. The purpose of the last breakout session of the workshop was to brainstorm ideas for new connection technologies for use in steel construction.

The ideas that resulted from the breakout sessions were then discussed among all the workshop participants. The primary feedback from the group centered on three needs. These included better production and fabrication processes to reduce tolerance requirements in the connector, the application of automated welding technologies - common in manufacturing - to steel erection, and machinery to eliminate or reduce the human involvement in the bolt-up process.

4.0 ANALYTICAL HIERARCHY PROCESS

The Analytical Hierarchy Process (AHP) is a multi-criteria (or multi-attribute) decision-making tool that was originally developed by Thomas L. Saaty in the early 1970's [11]. AHP is particularly useful when trying to rank alternatives based on the qualitative opinions of a group of experts. Since ranking the technologies and challenges that resulted from the workshop could not be carried out quantitatively without in-depth analyses, the opinions of the assembled steel industry experts were used with the AHP to rank the technologies and challenges.

The AHP is based on the pairwise comparison of the given alternatives taking only one criterion into consideration for each set of comparisons. Therefore, given n alternatives, the possible number of non-duplicative pairwise comparisons of the alternatives is

$$\sum_{i=1}^{n-1}(n-i)$$

If we are to rank these alternatives based on m criteria, then the above pairwise comparisons must be repeated m times (once for each criteria). For example, if n = 10 and m =5, then 225 comparisons would be required to complete the AHP! Hence, when using the AHP to rank several alternatives one must be careful to limit the number of alternatives and number of criteria in order to avoid an unwieldy number of comparisons. It was for this reason that the workshop participants were asked to choose only the top 5 technologies and challenges and only the top 3 criteria for each (resulting in a total of 60 comparisons).

The AHP also provides a means of checking the consistency of the pairwise comparisons so as to get a measure of the reliability of the data.

5.0 RESULTS

The results presented herein are limited to the results of the workshop's first two breakout sessions and the AHP analysis of those results.

5.1 Breakout Session One

The first breakout session resulted in a list of 12 technologies that the 17 workshop participants felt would be helpful in improving the productivity of steel construction. This session also resulted in a list of 10 criteria that the participants would use to rank the technologies. Tables 1 and 2 show the technologies and criteria, respectively, ranked in order of importance from top to bottom (based on the conglomeration of the participants' individual direct rankings).

5.2 Breakout Session Two

The second breakout session resulted in a list of 22 challenges that the 17 workshop participants felt the US steel construction industry currently faces. This session also resulted in a list of 9 criteria that the participants would use to rank the challenges. Tables 3 and 4 show the challenges and criteria, respectively, ranked in order of importance from top to bottom (based on the conglomeration of the participants' individual direct rankings).

5.3 AHP Results

In order to limit the number of pairwise comparisons conducted during the application of the AHP, only the top 5 technologies and challenges and the top 3 criteria were considered. For example, although in Table 2 the safety criterion ranked 4th in importance, it was not selected for the AHP analysis for the above reason. Ideally the AHP analysis would be conducted with all the technologies and challenges and all of their corresponding criteria.

Tables 5 and 6 show the results of the AHP analysis for the technologies and challenges, respectively. Table 7 shows the number of valid responses from which Tables 5 and 6 were generated. The valid responses were chosen based on an acceptable consistency ratio calculated as part of the AHP analysis [2].

Comparing Tables 1 and 3 with Tables 5 and 6, respectively, shows that the AHP results agree closely with the results of the workshop participants' manual ranking of the technologies and challenges. The AHP analysis also shows that apart from the "Material tracking" and "Material handling" technologies in Table 5 the final weighted scores in Tables 5 and 6 are not sufficiently different from one another to produce clear winners.

6.0 CONCLUSIONS AND FUTURE WORK

The workshop participants responded positively to the potential introduction of new technologies to the steel construction process and agreed that automation was needed in the industry. Many attendees volunteered to support future site visits and pilot studies. Based on the workshop response, the American Institute of Steel Construction is forming a steering committee to guide future research and the NIST Construction Metrology and Automation Group is making this research area a primary focus. A recommended research roadmap for automating steel construction will be presented in a forthcoming publication. This publication will include further analysis of the workshop results followed up by site visits and interviews with industry experts.

REFERENCES

1. Ricles, J., "Next Generation Steel Structures," Presentation delivered at the NIST/AISC Automated Steel Construction Workshop, Gaithersburg, MD, 6-7 June, 2002.
2. Haas, C. T., "Automation in Steel Erection," Presentation delivered at the NIST/AISC Automated Steel Construction Workshop, Gaithersburg, MD, 6-7 June, 2002.
3. Fleischman, R., Viscomi, B.V., and Lu, L.-W., "Development, analysis and experimentation of ATLSS connections for automated construction," *Proc., 1st World Conf. On Steel Struct.*, Acapulco, Mexico, 1992.
4. Albus, J. S., Bostelman, R. V., Dagalakis, N. G., "The NIST ROBOCRANE, A Robot Crane", *Journal of Robotic Systems*, July 1992.
5. Yamazaki, Y. and Maeda, J., "Construction process innovation through SMART system," *Proceedings of the 1st Congress on Computing in Civil Engineering*, Jun 20-22 1994, Washington, DC, USA, p 1583-1590.
6. Wakisaka, T., Furuya, N., Inoue, Y., and Shiokawa, T., "Automated construction system for high-rise reinforced concrete buildings,"., *Automation in Construction*, v9, n3, 2000, p 229-250.
7. Sakamoto, S. and Mitsuoka, H.; "Totally mechanized construction system for high-rise buildings (T-UP system)," *Proc., 11th International Symposium on Automation and Robotics in Construction*, May 24-26 1994, Brighton, UK.
8. Owen, F., Park, G., Haas, C., Gibson, G., and Traver, A., "Performance Testing of a Large Scale Manipulator to Determine Relative Utility of Several Operator Interfaces," *Proc., 14th International Symposium on Automation and Robotics in Construction*, Pittsburgh, PA, June 1997
9. Dunn, R., "Steel Erection and Challenges" Presentation delivered at the NIST/AISC Automated Steel Construction Workshop, Gaithersburg, MD, 6-7 June, 2002.
10. Ricles, J. and Schlafly, T., "Next Generation Steel Structures," White paper delivered at the NIST/AISC Automated Steel Construction Workshop, Gaithersburg, MD, 6-7 June, 2002.
11. Saaty, T. L., Multicriteria Decision Making: *The Analytical Hierarchy Process – Planning, Priority Setting, Resource Allocation*, 2nd Edition, RWS Publications, Pittsburgh, PA, 1990

Table 1. List of Helpful/Desired Technologies.

Ranking	Technology
Most	New connector technology
Desired	3D/4D CAD and data interchange
	Automated welding
	Material tracking technology (logistics)
	Piece movement technology (material handling)
	Plumbness technology
	Simpler method for installing and tensioning bolted connections
	Technology to locate components and objects
	Technology to create as-built models
↓	New steel technology
Least	Jack-up construction technology
Desired	Deck-sheet sidelap fastening technologies

Table 2. List of Criteria for Ranking the Technologies.

Ranking	Technology Criteria
Most	Cost savings
Important	Quality
↓	Speed/productivity
	Safety
	Minimization of rework
	Ease of Use
	Durability
	Time until 100% ROI
Least	Tolerance accommodation
Important	Make task attractive to labor

Table 3. List of Challenges Faced by the US Construction Industry.

Ranking	Challenges
Most	Reduce overall time to construct
Important	Reduce time from design to erection
	Need to optimize man-hours/ton
	Connection technology
	Efficient supply chain management from shop to erected state
	Facilitate information exchange
	Maximize efficiency of hoisting equipment
	Minimize cost of moment connections
	Reducing number of pieces (design stage)
	Standardize perimeter framing
	Inspection is labor intensive, time consuming, etc
	Minimize fall risk
	Confirming foundation accuracy
	Optimize bolting process on current connections
	Shop drawing time reduction (project critical path)
	Changing mind-sets of engineers, designers, constructors, etc.
	Accurate installation of base detail
	Streamline the code acceptance process
	Expand ability to use prefab modules
	Determining actual location of piece in lay-down area
	Ability to coordinate multiple cranes
	Improve quality control (tighten tolerances) of steel members
Least Important	Erection process susceptible to weather restrictions

Table 4. List of Criteria for Ranking the Challenges.

Ranking	Challenges Criteria
Most	Overall cost benefits
Important	Safety
	Productivity
	Quality
	Size of market
	Non-proprietary
	Code acceptance possibilities
Least	Time to market acceptance
Important	Durability/performance of end product

Table 5. AHP Results for the Top 5 Technologies and Top 3 Criteria.

Criteria (weights) / Technology.	Quality (0.41)	Cost Savings (0.34)	Productivity (0.25)	Final Weighted Score
Connectors	0.20	0.27	0.27	**0.24**
3D/4D CAD and data interchange	0.25	0.21	0.21	**0.23**
Automated welding	0.26	0.17	0.22	**0.22**
Material handling	0.15	0.18	0.15	**0.16**
Material tracking	0.14	0.17	0.15	**0.15**

Table 6. AHP Results for the Top 5 Challenges and Top 3 Criteria.

Criteria (weights) / Challenges	Overall Cost Benefits (0.41)	Productivity (0.35)	Safety (0.24)	Final Weighted Score
Reduce time from design to erection	0.30	0.19	0.15	**0.22**
Connection technology	0.20	0.18	0.30	**0.21**
Reduce overall time to construct	0.17	0.24	0.20	**0.20**
Need to optimize man-hours/ton	0.14	0.23	0.17	**0.18**
Efficient supply chain management	0.20	0.17	0.17	**0.18**

Table 7. Number of Valid Responses for Each Pairwise Comparison.

Comparison of:	Valid Responses
Pairs of challenge ranking criteria	7
Challenge pairs vs. Overall cost benefits	10
Challenge pairs vs. Productivity	9
Challenge pairs vs. Safety	13
Pairs of technology ranking criteria	8
Technology pairs vs. Cost savings	14
Technology pairs vs. Quality	12
Technology pairs vs. Productivity	14

Error Modeling for Automated Construction Equipment

by

Yong K. Cho*, Carl T. Haas**, S.V. Sreenivasan***, and Katherine Liapi****

ABSTRACT: This paper introduces the statistical error modeling approach for a computer-controlled large-scale manipulator (LSM). The LSM is sufficiently representative of several types of construction equipment to be able to serve as a general test bed. In the analysis, three factors which are measurable in real time: distance, hydraulic pressure, and payload, were varied to determine their influence on position errors in the LSM. It was shown that with an integrated multi-variable regression model, about 30% of the mean positioning error of the LSM can be reduced without the mid of fixed external reference systems.

KEYWORDS: automation; kinematics; manipulator; position error; regression analysis

1. INTRODUCTION

Hydraulically actuated construction equipment is rapidly being retrofitted with robotic control capabilities by several major manufacturers such as dozers, graders, and excavators equipped with various position sensors reflects a movement in the construction industry towards improving productivity, efficiency, and safety [6].

The reduction of the cumulative position error caused by backlash, deflection, sensor error and other factors, however, still remains as one of the key issues in operating autonomous or semi-autonomous construction equipment. Further, large scale construction equipment, particularly the ones actuated hydraulically, possess lower end-point positioning accuracy than electrically actuated robotic devices used in controlled environments.

Understanding the causes and propagation of errors in automated construction equipment is very important for precise operation in a field environment. Large manipulators such as forklifts, excavators, and the LSM pose particularly difficult equipment control problems. Much of the manipulator error is due to accumulated feedback errors of the rotary and prismatic joint sensors, and lost motion due to backlash. Working conditions such as variations in hydraulic pressure supply and presence of large payloads influence the error attributes as well. Kinematic and dynamic states are also an influence.

Most errors arising from the operation of the manipulator are considered random errors which are unavoidable and are usually reduced through stochastic experimentation [1]. Even if it is theoretically possible to derive a mechanistic mode of all sources and their interdependence, it is computationally untenable for real time manipulator control, and sensor requirements for independent input variables are currently unfeasible in a construction environment. Position error of an economically sensor equipped manipulator can be reduced by providing the correction factors resulting from a statistical analysis of on-hand samples. The results can be organized into statistical equations with correction factors for real-time equipment

* Corresponding Author, Assistant Professor, Dept. of Industrial Studies, University of Wisconsin, Platteville, WI 53818. Tel.: +1-608-342-1604 ; fax: +1-608-342-1254. E-mail: choyo@uwplatt.edu
** Professor, Dept. of Civil Engineering, The University of Texas, Austin, Texas 78712
*** Associate Professor, Dept. of Mechanical Engineering, The University of Texas, Austin, Texas 78712
**** Assistant Professor, Dept. of Civil Engineering, The University of Texas, Austin, Texas 78712.

feedback control, much like those used for artillery fire control.

2. OBJECTIVE AND METHODOLOGY

The main objective of the research presented here is to improve automated manipulator positioning capability based on regression modeling of manipulator position errors and working conditions.

The methodology pursued includes the following basic steps:

1) Select a set of well distributed known positions in a manipulator's working space,
2) Command a robotically controlled manipulator's end-point to the selected positions,
3) Measure actual positions of the resulting end-point, and calculating position errors,
4) Repeat the above steps for varied working conditions as defined by the following causal factors: distance, hydraulic pressure, and payload,
5) Analyze the results with the regression modeling method, and
6) Develop an algorithm based on the regression modeling and apply the algorithm to adjust the commanded positions to reduce errors.

3. KINEMATICS

Inverse and forward kinematics equations, computer algorithms, feedback encoders, control interfaces were implemented from the previous research efforts for the various control strategies for the LSM [4]. An illustration of the large scale manipulator is shown in Fig. 1. As seen in the figure, the LSM has swing, lift, telescope, rotate, roll, and pivot joints in total 6 degrees of freedom (DOF).

The LSM control system calculates a series of joint angles through which to move the joints in order for the end-effector frame to move from its initial location to its final location. Then, the link transformations can be multiplied together to find the single transformation that relates a frame to another frame. As with vectors and rotation matrices, a symbol T is called a transformation operator [3]. Here, ${}^{Base}_{Pivot}T$ describes Pivot frame relative to Base frame and forms the following transform equation:

$$ {}^{Base}_{Pivot}T = {}^{Base}_{Swing}T \; {}^{Swing}_{Lift}T \; {}^{Lift}_{Tele}T \; {}^{Tele}_{Rotate}T \; {}^{Rotate}_{Roll}T \; {}^{Roll}_{Pivot}T $$

Each joint angle can be computed from the transform equation.

4. ACCURACY TESTS

Accuracy may be defined as the magnitude of the difference between the desired value and actual value of a measurement. In this study, accuracy is measured relative to the manipulator's base coordinate system (see Fig. 2).

4.1 Sample Regression Analysis

There are many existing random error sources affecting the final position of the manipulator's end-effector. Thus, in this study, to predict and to adjust for the position errors from the uncertain random error sources, the sample regression modeling method was selected to determine whether or not a relationship exists between error variables and to predict the nature of the relationship of correlated random variables.

To ensure a consistent test process, six different kinematic states of the LSM were chosen as known target positions (Table 1). They were chosen to represent the full range of configuration that was anticipated to affect the various random error sources under the different conditions of the chosen three independent variables: distance, hydraulic pressure, and payload.

4.1.1 Accuracy with respect to Distance

Errors with respect to the distance variable were analyzed under several different conditions: five different payloads and five different hydraulic pressures in the six fixed positions.

In this study, the distance is defined as the Cartesian positional difference between a base point (0,0,0) of the LSM's test bed and the kinematically calculated position (x, y, z) of the end-effector.

The accuracy test result for individual axial accuracies with signed values were tested by distances ranging from 0 to 4.5 m. The desired position was calculated from the kinematics equations, and the actual position was manually measured. Then, the actual positions of the end-effector were compared to the kinematically calculated positions. Then the directional (axial) error attributes with respect to the distance variable were analyzed with collected sample data (see Table 2). Here, n indicates the sample size (5 pressures × 6 positions + 5 payloads × 6 positions = 60 samples).

While distance increased, as results, the directional errors in the Z axis increased, which relates to or is explained by the fact that angular joint error increases with respect to distance. Correlation analysis might be useful to determine the deflection in z coordinates according to the directional distance in X and Y axes ($\sqrt{(x^2+y^2)}$) which describes how far the arm is cantilevered out.

4.1.2 Accuracy with respect to Hydraulic Pressure

The current LSM's hydraulic system configuration is a single power source that feeds a parallel network of actuators. with a single branch for each joint. The computer control signal determines the strength of the current to the valve solenoids and thus the opening of the control valve. This determines the flow rate of fluid into the actuator. According to the LSM payload and location in workspace, the control signal strength changes and thus the flow rate. To keep the joint moving at a steady speed under an increasing load, the valve opening increases as the load increases. This can be done by the pressure compensation valve. So even with an increasing load, the LSM's pressure compensation valve keeps the actuator velocity steady [5].

Even with the advanced hydraulic flow control system. the LSM still exhibits more stiction. slop, and backlash in the joints and hydraulic actuators than conventional industrial robots which are small, fast. electrically actuated. Besides, the LSM can be manufactured to lower tolerances than can be the industrial robots. Thus, as another variable, hydraulic oil pressure was selected to be examined in order to find the error attributes of the LSM. Due to maintenance concerns, hydraulic pressure was limited to a maximum of 10345 kPa (1500 psi).

LSM Table 3 shows the directional error attributes with respect to the hydraulic pressure. All the linear regression equations indicate the error decreases toward zero while the hydraulic pressure increases.

Similar to the previous accuracy test toward distance, hydraulic pressure showed some relationship with the directional error in the Z axis. The directional error in the Z axis decreased while hydraulic pressure increased. The increased hydraulic strength increases the joint speed, which might ultimately reduce the error caused from stiction and load. In a sense, the system is stiffer as well.

4.1.3 Accuracy with respect to Payload

As another variable, payload was examined in order to find the error attributes of the LSM. Payload affects the inertia and speed of the manipulator, which can yield an overshoot error and more deflection. Four different pipes were used for this test. The weight of the pipes ranged from 40 kg to 386 kg.

To ensure a consistent test process, the experiment was performed under 8276 kPa (1200 psi) of hydraulic pressure. Each pipe was lifted into the chosen kinematic states of the LSM. Fig. 3 shows one of the chosen kinematic states.

The directional (axial) error attributes with respect to the payload variable were analyzed and summarized in Table 4.

The coefficient of determination r^2 for the X axis (0.507) and the Z axis (0.531) indicate that the payload significantly affects the accuracy of the X axis and the Z axis. It is clear that payload accentuates the deflection which explains the increased error in Z coordinates. In addition, payload makes the LSM move sluggish which might cause in more errors toward the X axis which has a 4.5 m range than toward the Y axis which has 2.2 m range. Payload makes the lifting and telescoping motions more sluggish, which might ultimately cause some systematic relationship with errors in x and z coordinates. Also, x and z coordinates relate to degree of cantilever and therefore deflection.

Although there was almost no relationship between payload and the directional error in the Y axis, the systematically increased conditional variance of accuracy in the Y axis might be caused from erratically applied error sources such as stiction and overshoot to the swing joint.

4.2 Error Adjustment with Multi-variable Regression Models

The multi-variable regression analysis of the LSM's position with respect to distance, hydraulic oil pressure, and payload was modeled to adjust the error of each axis of the end-effector as follows:

$$\Delta X=-1.229+0.182D_{istance} + 0.0000528P_{ressure} - 0.00264\, P_{ayload},\ (r^2=0.262)$$

$$\Delta Y=0.627-0.1D_{istance} - 0.000018P_{ressure} + 0.0000988P_{ayload},\ (r^2=0.028)$$

$$\Delta Z= -0.884 - 0.141\, D_{istance} + 0.0001075\, P_{ressure} - 0.00397\, P_{ayload},\ (r^2=0.483)$$

Here, ΔX, ΔY, and ΔZ indicate the regression of the multi-variable model on the LSM's axial position error

By applying the three multi-variable equations to the same data set obtained from the LSM's position accuracy test, given the three conditional variables, there were some error reductions (30.7%) in the overall accuracy shown in Table 5. In a theory, the substantial number of the unsolved errors (69.3%) may be statistically solved by adding more conditional variables.

While the average errors of the X axis and the Z axis were reduced by 18.88% and 50.55% respectively, the average error of the Y axis somewhat increased by using the multi-variable equations. This means that the multi-variable equation (ΔY) does not explain the Y axis errors very well. Also, it can be inferred from the equation in which r^2 is 0.028, that the multiple variables have little effect on the Y axis accuracy. Therefore, while adjusting the X and Z axis values based on Equation ΔX and Equation ΔZ, the kinematically calculated original Y axis value was not adjusted.

4.3 Final Performance Test

The obtained three equations were then applied to the original forward kinematics equations to adjust for the original x-y-z coordinates. Then, the inverse kinematics provides adjusted six joints angles based on the adjusted position data. To verify the error attributes found from the error modeling tests, several material placement tests were conducted on the LSM's test bed. This paper introduces one of the tests here.

A stylus was attached to the test load (386 kg) in the LSM's jaws pointing toward a target (see Fig.4). By using a developed laser rangefinder system and string encoders [2], a Cartesian coordinates of a center of cross hairs on a target plane was measured (325.63cm, 10.77cm, 77.09cm). The purpose of this test was to compare the LSM's accuracy when its error was adjusted and unadjusted under a certain working condition (with 3.348m distance, a 386 kg test load, and a 8276 kPa (1200 psi) hydraulic pressure supply). A test result shows that the LSM has better accuracy when a commanded position was modified based on the developed error modeling as follows:

a) Error without adjustment: 4.29 cm
b) Error with adjustment: 1.97 cm

Here, the errors were measured by the distance between the center of the cross hairs and the end of the stylus.

5. CONCLUSION

The error attributes of a hydraulically actuated large scale manipulator were analyzed by using regression analysis. In the regression analysis, three variables, distance, hydraulic pressure, and payload, were individually varied to find the position error attributes of the LSM. Distance had a somewhat significant effect on the directional error in the Z axis (as distance increased, random errors in the Z axis increased). Hydraulic pressure and payload had significant effect on the overall error (as hydraulic pressure decreased and payload increased, random error increased). Although, the testing was performed in a small working volume due to the limited mobility of the LSM on its fixed frame, this study reduced about 30% of the average positioning errors of the LSM with an integrated multi-variable regression model. Thus, it is sufficient to indicate whether a descriptive model or a regression model is feasible. The substantial number of the unsolved errors (about 70%) may be statistically solved by adding more conditional variables which possibly affect the manipulator's position errors.

6. REFERENCES

[1] Barry, B.A. (1978). *Errors in Practical Measurement in Science, Engineering, and Technology.* John Wiley & Sons, Inc., New York.

[2] Cho, Y., Haas, C., Sreenivasan, S., and Liapi, K. (2002). "A framework for rapid local area modeling for construction automation." Journal of Automation in Construction, 11(6), pp. 629-641.

[3] Craig, J. (1986). *Introduction to Robotics: Mechanics & Control.* Addison-Wesley.

[4] LeBlond, D., Owen, F., Gibson, G., Haas, C., and Traver, A.(1998). "Performance Testing and Characterization of Advanced Construction Equipment." ASCE Journal of Construction Engineering and Management. 124(4), pp. 289-296.

[5] Owen, F.C. (1998). "Development of Smooth Resolved Motion and Computer- Assisted Teleoperation for a Large Scale Manipulator." *Ph.D. Dissertation*, Dept. of Mechanical Engineering, The University of Texas at Austin.

[6] Singh, S. (1997). "The State of the Art in Automation of Earthmoving." ASCE Journal of Aerospace Engineering, Vol 10 (4), October.

Table 1. Chosen Kinematic States

Kinematic States					Calculated Global Position of End-effector (cm)		
Swing (°)		Lift (°)	Telescope (cm)	Rotate (°)	X	Y	Z
(1)	-20	11	0	20	157.9992	-54.0685	73.5178
(2)	5	50	30.48	0	364.4036	31.8821	186.7875
(3)	-10	36	60.96	0	321.0855	-56.6166	95.7560
(4)	-2	31	16.76	-10	276.6731	-11.3020	112.2213
(5)	10	21	0	0	216.2586	38.1335	95.5121
(6)	0	54	48.77	0	393.9296	0	197.3946

(The values of Roll and Pivot joints are 0°)

Table 2. Directional (Axial) Error Attributes with respect to Distance

Axis	**Regression Model**	**n**	r^2
X	$\Delta x= -0.962+0.180\ D_{istance}$	60	0.059
Y	$\Delta y=0.493-0.1\ D_{istance}$	60	0.026
Z	$\Delta z=0.0318-0.2411\ D_{istance}$	60	0.122

Table 3. Directional (Axial) Error Attributes with respect to the Hydraulic Pressure

Axis	**Regression Model**	**n**	r^2
X	$\Delta x=-0.58+0.00001412\ P_{ressure}$	30	0.003
Y	$\Delta y=0.549-0.000043\ P_{ressure}$	30	0.033
Z	$\Delta z=-1.29+0.00008942\ P_{ressure}$	30	0.178

Table 4. Directional (Axial) Error Attributes with respect to the Payload

Axis	**Regression Model**	**n**	r^2
X	$\Delta x=0.151-0.00397\ P_{ayload}$	30	0.507
Y	$\Delta y=0.124+0.0002431\ P_{ayload}$	30	0.003
Z	$\Delta z=-0.267-0.00459\ P_{ayload}$	30	0.531

Table 5. Axial Error Adjustment for Test Data Set

Average Accuracy	X	Y	Z	Total Accuracy
Before Adjusted (cm)	0.5731	0.3859	0.7309	1.1717
After Adjusted (cm)	0.4649	0.4009	0.3614	0.8115
% Reduced	18.88	-3.66	50.55	30.74

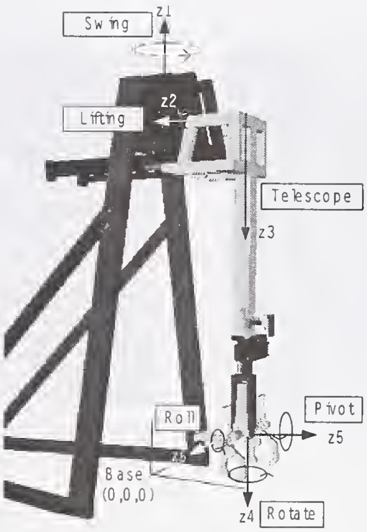


Fig. 1. 6 DOF Kinematic Configuration for the LSM

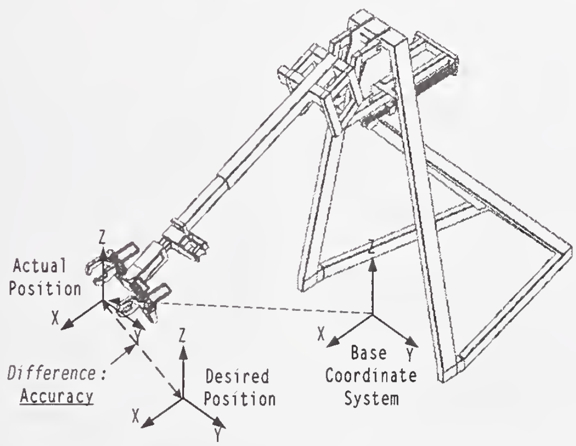


Fig.2. Illustration of the Accuracy Method relative to the Base Coordinate System

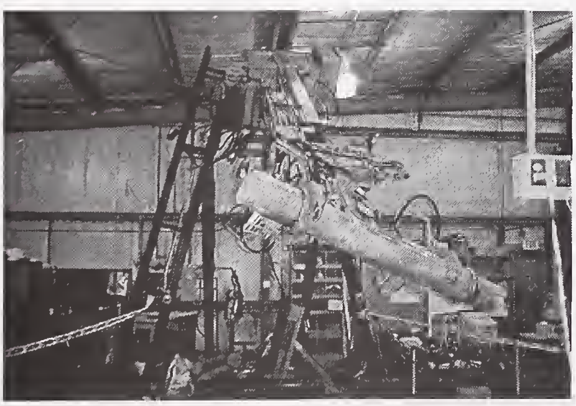


Fig. 3. One of the chosen Kinematic States

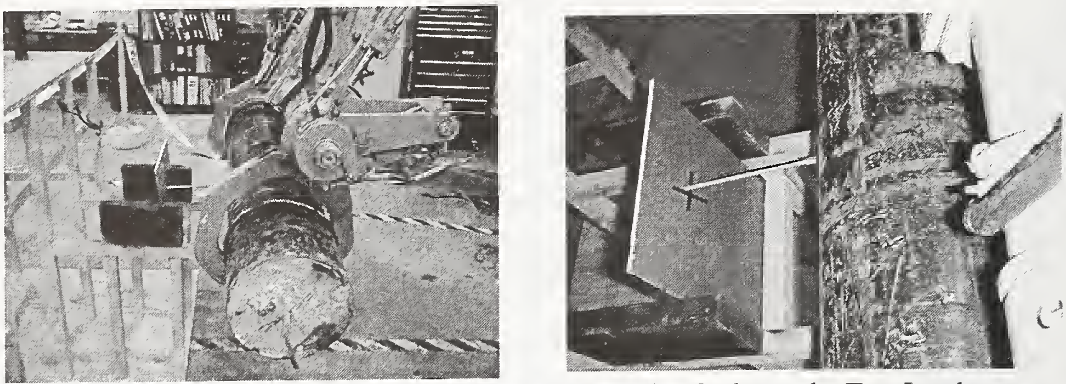


Fig. 4. The Accuracy Test was conducted with a Stylus and a Test Load

Development of a Robotic Structural Steel Placement System [1]

by

Alan M. Lytle[2], Kamel S. Saidi[2], and William C. Stone[2]

ABSTRACT: The NIST Construction Metrology and Automation Group, in cooperation with the NIST Intelligent Systems Division, is researching robotic structural steel placement as part of a project to develop an Automated Steel Construction Testbed. This project was initiated in response to industry requests for advanced tools and methodologies to speed the erection of steel structures while maintaining or enhancing standards for worker safety and facility reliability. This initial effort integrates and extends prior NIST research in robotic crane employment, tele-operated steel beam placement, laser-based site metrology, construction component tracking, and web-enabled 3-D visualization.

KEYWORDS: construction automation, path planning, robotics, VRML, 3-D coordinate measurement systems

1.0 INTRODUCTION

The American Institute of Steel Construction has expressed a need for a 25 % reduction in time to erect steel structures. In response to this request, the NIST Construction Metrology and Automation Group (CMAG) is developing a robotic structural steel placement system for the testing and validation of advanced tools, methodologies, and standards for automated steel construction. This effort, the first phase of CMAG's Automated Steel Construction Testbed (ASCT) project, extends prior NIST research in construction component tracking [1] and real-time construction metrology [2]. The work also extends prior efforts in tele-operated steel beam placement [3] to include autonomous control of a six degree-of-freedom (DOF) robotic crane.

The base platform is the NIST RoboCrane, which is an inverted Stewart platform parallel link manipulator [4]. RoboCrane's Real-Time Control System (RCS) is augmented with absolute cartesian position feedback from a laser-based site measurement system (SMS) for trajectory planning and dynamic control. Robot and construction component position data is displayed in a virtual world model of the steel placement operation for supervisory review and control.

The steel structure to be assembled consists of a beam and two columns connected with the ATLSS[3] quick connector. The ATLSS, developed at Lehigh University [5], was chosen because it requires no bolting or welding.

This paper will discuss the concept, current development and future implementation of the robotic structural steel placement system.

2.0 OPERATIONAL CONCEPT

Four laser transmitters are positioned on the site perimeter to illuminate the work volume of RoboCrane with reference beams (Figure 1). A field worker using the SMS digitizing wand then

---

[1] Official contribution of the National Institute of Standards and Technology (NIST); not subject to copyright in the United States.

[2] NIST, BFRL, Construction Metrology and Automation Group, Mail Stop 8611, Gaithersburg, MD 20899-8611; alan.lytle@nist.gov, kamel.saidi@nist.gov, william.stone@nist.gov.

[3] Certain commercial equipment, instruments, or materials are identified in this report in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

creates a digital model of the construction plane including any obstacles.

This world model is then updated with the positions of the as-built structure and the target beam by the field worker through a process of automatic part identification (barcode), part model database access, and part fiducial point measurement (Figure 2). The current pose of RoboCrane is measured from onboard SMS sensors, and the path planner calculates the required transformations for beam pickup and delivery. RoboCrane then executes the movements (Figure 3). The world model, including dynamic tracking of RoboCrane, is displayed to an operator through a visualization system based on the Virtual Reality Modeling Language (VRML).

3.0 SYSTEM OVERVIEW

The current configuration of the Automated Steel Construction Testbed has five primary components. These include:

- RoboCrane
- Site Measurement System
- Component Tracking System
- High-level ASCT Controller
- Visualization System

2.1 RoboCrane

RoboCrane is an innovative, cable-driven, manipulator invented by the NIST Intelligent Systems Division (ISD) and further developed and adapted for specialized applications over a period of several years [6,7,8]. The basic RoboCrane is an inverted Stewart platform parallel-link manipulator with cables and winches serving as the links and actuators, respectively. The moveable platform, or "lower triangle," is kinematically constrained by maintaining tension in all six cables that terminate in pairs at the vertices of an "upper triangle" formed by the cable support points. This arrangement provides improved load stability over traditional lift systems and enables 6 DOF payload control.

The version of RoboCrane used in this project is the Tetrahedral Robotic Apparatus (TETRA). In the TETRA configuration, all winches, amplifiers, and motor controllers are located on the moveable platform. The upper triangle only provides the three tie points for the TETRA cables, allowing the device to be retrofitted to existing overhead lift mechanisms.

2.2 Site Measurement System

The SMS uses commercially available positioning technology (*3D-I*) produced by Arc Second, Inc. in the *Constellation* and *Vulcan* product families[3]. (*3D-I*, *Constellation* and *Vulcan* are registered trademarks of Arc Second, Inc.) These systems use stationary, active-beacon laser transmitters and mobile receivers to provide millimeter-level position data.

2.2.1 SMS Description

Both *Constellation* and *Vulcan* systems use eye-safe laser transmitters to triangulate the position of a tuned optical detector. Each transmitter emits two rotating, fanned laser beams and a timing pulse. Elevation is calculated from the time difference between fan strikes. Azimuth is referenced from the timing pulse. The field of view of each transmitter is approximately 290° in azimuth and +/- 30° in elevation/declination. The recommended minimum and maximum operating ranges from each transmitter are 5 m and 50 m, respectively.

Line-of-sight to at least two transmitters must be maintained to calculate position. The *Constellation* receivers each track up to four transmitters and wirelessly transmit timing information to a base computer for position calculation. The *Vulcan* system is a self-contained digitizing tool with two optical receivers on a rigid pole. A vector projection along the line formed by the two optical detectors allows 3-D measurement of the tool tip. *Vulcan* can track only two transmitters at one time; however, the transmitter selection can be manually switched between any of the four available. Recovery of positional data following momentary signal blockage takes approximately one second.

2.2.2 Prior *3D-I* / Mobile Robot Integration

Early efforts to use *3D-I* laser technology for mobile robot navigation showed that although the system was capable of guiding a mobile robot [9], its use was restricted due to loss of track at relatively low vehicle speeds [10]. Upgrades to the positioning technology continued and a successful combination of indoor 2-D map creation and autonomous navigation was demonstrated in a research project at the Rochester Institute of Technology [11]. Subsequently, a single receiver *Constellation* system was installed on an autonomous lawn mower at the Carnegie Mellon University Field Robotics Center and provided positional reference in a large outdoor setting [12].

2.2.3 The SMS on RoboCrane

Three SMS receivers are mounted on RoboCrane at the vertices of the lower triangle (see Figure 4). The receiver locations are registered to the manipulator during the initial setup process in the local SMS coordinate frame. For convenience, all measurements are calculated in the local SMS coordinate frame, though if required, mapping to an existing world coordinate frame could be accomplished. Receiver timing signals and diagnostic data are wirelessly transmitted to a base station computer running Arc Second's proprietary position calculation software. Position and SMS diagnostic information are polled at approximately 7 Hz using a NIST-developed data communications application. Position data from the three receivers is used to calculate RoboCrane's pose (position and orientation). Diagnostic data such as number of visible transmitters, excess signal noise or multipath reflections is also provided for each position calculation and is used to assess the quality of individual position fixes.

2.3 Component Tracking System

The NIST Construction Metrology and Automation Group has conducted research in construction component tracking for several years. NIST developed the "Comp-TRAK" system, which combines field-wearable computing, automatic identification (barcode and RFID), laser metrology, wireless database access, and VRML-based 3-D model visualization to enable tracking of discreet components on a construction site. The feasibility of such a system was shown during a study tracking delivery and final placement of structural steel members on a NIST construction project [13]. During this study, the Comp-TRAK system was successfully used to identify components, remotely access component library data, guide the field user in measuring current part location, and remotely update a project management database with part data including location within the site coordinate frame.

Elements of Comp-TRAK will be used in the robotic structural steel placement system to locate randomly-placed target components and update the world model through the following sequence:

(a) The field user identifies the component.
(b) The component library is accessed and a model is displayed with specified measurement points (fiducial points).
(c) The field user measures observable fiducial points with the SMS digitizing wand.
(d) The component position and orientation is calculated and transferred to the world model.

2.4 High-level ASCT Controller

A high-level ASCT controller will provide the following task and data management functions for the overall system:

- Goal state management
- World model maintenance
- Navigation
- RoboCrane controller interface

The master goal of "pick and place steel beam" is divided into various sub-goals by the ASCT controller. These sub-goals define various states RoboCrane must achieve to execute the overall task. Examples include (1) maneuver to proper gripper orientation, (2) lower gripper onto beam, (3) grasp beam, (4) maneuver to pre-dock position, and (5) dock beam.

The existence and locations of components of interest are maintained in the ASCT world model. Data to update the world model are provided either by the component tracking system or by tracking RoboCrane's pose when the crane is carrying an object.

RoboCrane's pose within this world model will be estimated through Kalman filtering of both the SMS position fixes and the RCS position estimates derived from winch encoder feedback. Based on world model component locations, RoboCrane's pose, and volume operating limits, a series of waypoints will be calculated for each goal state. The required platform transformations for the state changes and corresponding cartesian velocity commands (Translation: x, y, z; Rotation: roll, pitch, yaw) are calculated and sent to the RoboCrane controller via a communications interface. The RoboCrane controller, a version of the NIST Real-time Control System implemented by ATR (Advanced Technology & Research Corporation)[3], then converts the cartesian velocity commands to winch controller input to execute the desired movement. Closed-loop position feedback from the SMS position fix enables periodic modification of the path by the ASCT controller until the desired goal state is reached.

2.5 Visualization System

A VRML-based 3-D visualization system will also be used to provide remote visual feedback to an operator or supervisor. Elements of the world model – the construction plane, RoboCrane, and the target components – will be modeled in VRML 97 [14] and displayed within a browser environment giving an observer 3-D "fly-through" review capability. A socket connection between the ASCT controller and the VRML environment will provide pose updates for RoboCrane, components being moved, and other elements of interest. The VRML object models within the visualization system will then be repositioned by a Java applet using the External Authoring Interface browser extension. This provides a non-proprietary, open standard, low-bandwidth method of displaying a 3-D representation of robot operations within the work site.

3.0 CONCLUSIONS

This project will demonstrate autonomous steel structure assembly (pick and place) using a robotic crane and a laser-based SMS. A digital model of the work site is created with the SMS and then the same measurement system is used for closed loop feedback to precisely control the pose of the steel components during assembly.

4.0 FUTURE WORK

Although this work will demonstrate the ability to perform autonomous steel beam pick and place, it will do so in a fairly structured environment that remains static after initial measurement. The measurement process itself, although simple, still requires a human operator within the site to provide the initial digital model. There are currently no sensors, external or on-board RoboCrane, which would enable any reaction to dynamic changes within the work site without human intervention.

CMAG is currently researching automatic scene meshing and object recognition using high-resolution LADAR (laser detection and ranging) systems. In future work, LADAR and/or standard optical imaging systems will be used to develop and maintain the world model as well as provide obstacle avoidance and docking support. The use of the SMS technology to provide autonomous control of cable suspended robots will also be studied for other applications such as aircraft maintenance and shipbuilding. As additional sensing systems are employed, the SMS technology will also be used to study the performance metrics of other tracking technologies. The ASCT treats sensor input in a modular fashion; thus, future versions could use other positioning technologies such as phase differential GPS in certain applications to replace the present laser-based SMS.

5.0 REFERENCES

1. Furlani, K. M. and Pfeffer, L. E., "Automated Tracking of Structural Steel Members at the Construction Site." *Proc., 17th International Symposium of Automation and Robotics in Construction*, Taipei, Taiwan, Sept. 18-20, 2000.
2. Stone, W. C, Pfeffer, L. E. and Furlani, K. M., "Automated Part Tracking on the Construction Job Site." *Proc., 4th ASCE Specialty Conference and Exposition/ Demonstration on Robotics for Challenging Situations and Environments*, pp. 96-103, Albuquerque, New Mexico, Feb. 27 – Mar. 2, 2000.
3. Stone, W. C., Reed, K. A., Chang P. C., Pfeffer, L. E. and Jacoff, A. S., "NIST Research Toward Construction Site Integration and Automation." Journal of Aerospace Engineering, Apr. 1999, pp. 50-57.
4. Stewart, D., "A Platform with Six Degrees of Freedom," Proc. of the Inst. of Mechanical Engineering, Volume 180(15), Part I:371-386, 1965-1966.
5. Fleischman, R., Viscomi, B.V., and Lu, L.-W., "Development, analysis and experimentation of ATLSS connections for automated construction." *Proc., 1st World Conf. On Steel Struct.*, Acapulco, Mexico, 1992.
6. Albus, J. S., Bostelman, R. V., Dagalakis, N. G., "The NIST ROBOCRANE, A Robot Crane", Journal of Robotic Systems, July 1992.
7. Bostelman, R., Albus, J., Dagalakis, N., Jacoff, A., "RoboCrane Project: An Advanced Concept for Large Scale Manufacturing," *Proc, Association for Unmanned Vehicles Systems International*, Orlando, FL, July 1996.
8. Bostelman, R. , Albus, J., Stone, W., "Toward Next-Generation Construction Machines." *Proc. Of American Nuclear Society 9th International Topical Meeting on Robotics and Remote Systems*, Seattle, WA, March 4-8, 2001.
9. Beliveau, Y. J., Fithian, J. E., and Deisenroth, M. P., "Autonomous Vehicle Navigation with Real-Time 3-D Laser Based Positioning for Construction." Automation in Construction, (5) 1996.
10. Lytle, A. M., "Development of a Remotely Piloted Vehicle (RPV) with Real-Time Position Measurement (RtPM) for Hazardous Waste Site Characterization." Master's Thesis, Virginia Tech, Blacksburg, VA, 1994.
11. Rochester Institute of Technology, Society of Manufacturing Engineers (Student Chapter 18), Robot Construction Project, 2001. www.rit.edu/~smewww/project/rc/rc.html
12. Batavia P. H., Roth S. A., Singh S., "Autonomous Coverage Operations In Semi-Structured Outdoor Environments." *Proc., IROS*, Lausanne, Oct. 2002.
13. Furlani, K. M., Latimer, D. T., Gilsinn, D. and Lytle, A. M., "Prototype Implementation of an Automated Structural Steel Tracking System." *Proc., 19th International Symposium on Automation and Robotics in Construction*, Gaithersburg, MD, Sept. 23-25, 2002.
14. ISO/IEC 1477-1, "The Virtual Reality Modeling Language (VRML)." 1997, www.web3d.org

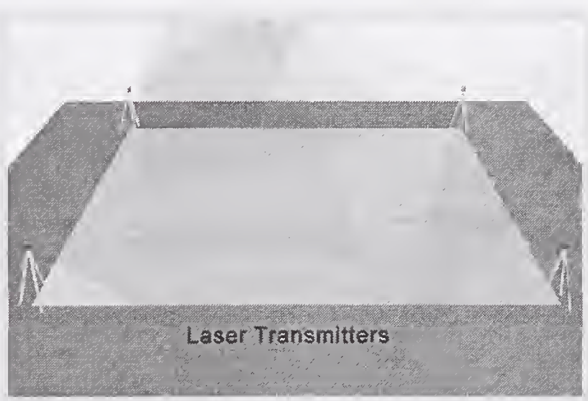


Figure 1: Graphic denoting the illumination of the work site with the SMS.

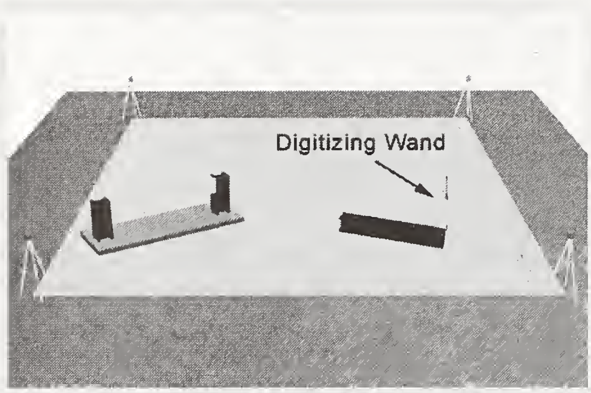


Figure 2: Graphic denoting measurement of component locations with the SMS digitizing wand.

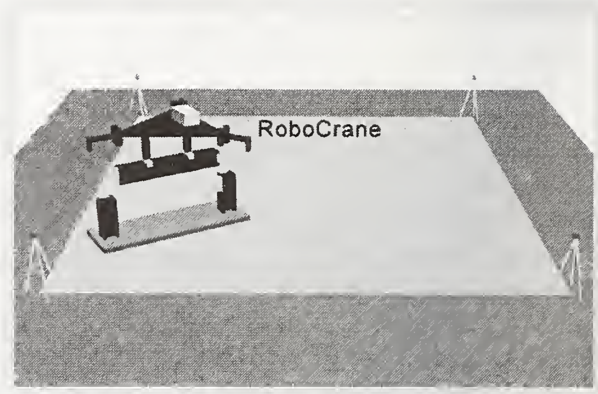


Figure 3: Graphic denoting steel beam placement with RoboCrane.

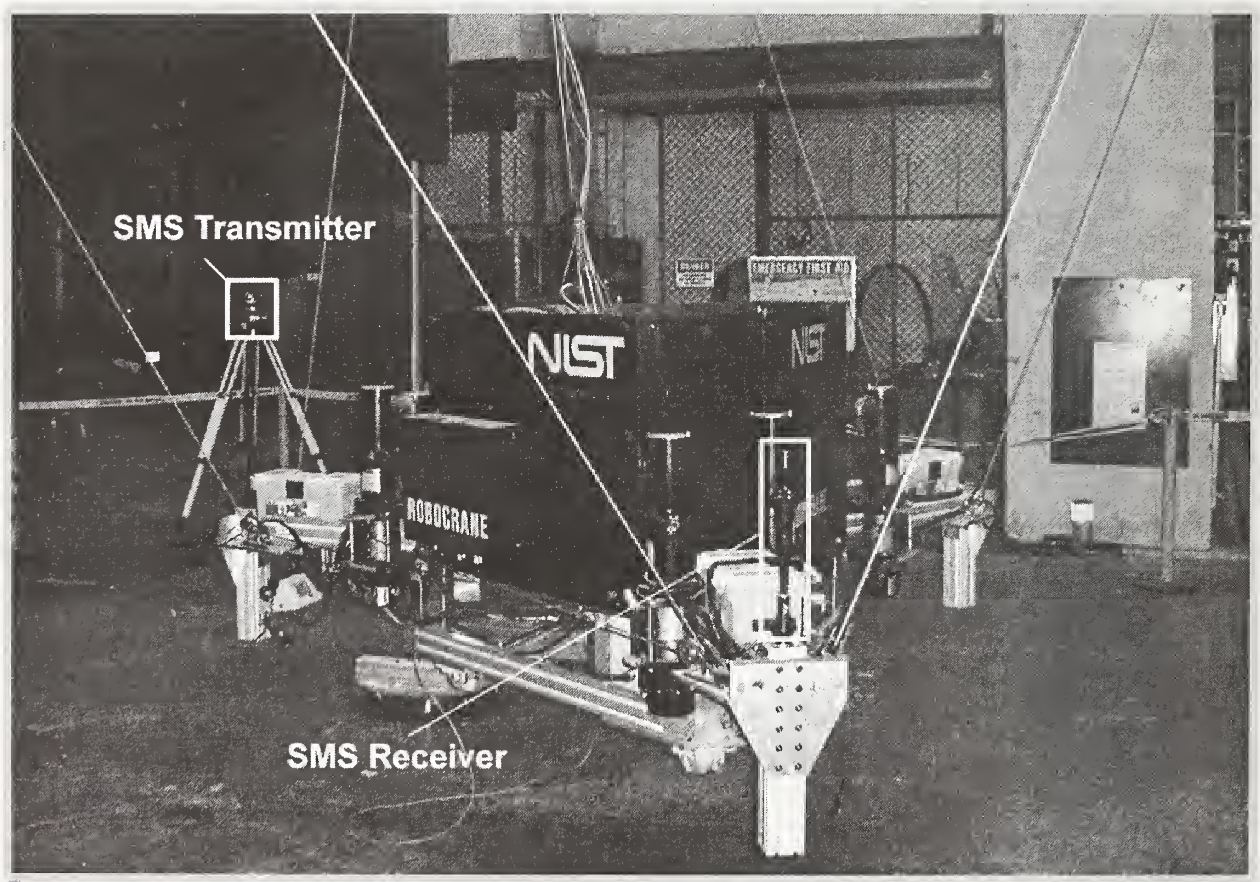


Figure 4: Photograph of RoboCrane with the SMS.

ROBOTS FOR SPACING OF WOOD

Thomas Bock*, Dimitry Parshin, Tatjana Souetina, Aleksei Boulgakov*****

**Technical University Munich, bulgakow@bri.arch.tu-muenchen.de*
*** Rostov State Building University, RUSSIA, pragma_plus@mail.ru*
**** Moscow Building University, t_a_souetina@mail.ru*

Abstract: Construction is of great interest for using of woodprocessing robotics. The structure, kinematic and dynamic models of robots for material location and spacing as well as woodprocessing have been considered in the paper. Taking into account specific features and dynamic characteristics of robot, the recommendations on movements planning and forming laws of control have been given. In conclusion recommendations for obtaining truthful measuring data and on the composition of robot software have been provided.

Keywords: Motion control; robot kinematics; wood processing.

1. INTRODUCTION

When producing elements of prefabricated wooden houses a great deal of work on location and spacing of wood and shaping of holes, recesses, decorative patterns is done. The projects individual peculiarities require the application of technological equipment with quick readjustment and opportunity to prepare control programs on the model during a short period of time. While solving this problem much attention is paid to the robotization of the mentioned above operations and creating robotic systems. The successful solution of robotization tasks is first connected with the development of original kinematic structures and competent structural analysis. One more significant problem of robotization of wood processing operations is setting the trajectory for cutting tool movement and provision of its purposeful movement along these trajectories with definite orientation.

2. ROBOT STRUCTURE, KINEMATICS AND DYNAMIC MODELS

The analysis of jobs connected with location and spacing of wood as well as shaping of holes and decorative patterns has shown that in the basis of the robotic system there should be a rectangular 3-coordinate gantry robot (fig.1). This cell provides the working tool movement in the plane of a working table as well as lifting, lowering and pressing of the tool. The second cell is an orienting working head providing changes in the tool position relative to the working plane or its rotation around the axis Z. The main kinematics relations defining the nature of the wood processing robot motions are presented by the system of the form

$$x(t) = q_1(t) + l\sin(q_4(t))\sin(q_5(t));$$
$$y(t) = q_2(t) + l\cos(q_4(t))\sin(q_5(t));$$
$$z(t) = q_3(t) + l\cos(q_5(t)),$$

where $q_i(t)$ are generalized robot's coordinates, l is the length of the end link.

The laws of changing the tool phase coordinates $x(t), y(t), z(t)$ and its orientations $\theta(t), \gamma(t)$ are determined by shape and view of the pattern being fulfilled. The angle θ specifies the tool pitch relative to the plane XY, and the angle γ - the direction of the inclination being read from the axis X. The values of these parameters and the laws of their change in time are formed at the stage of planning robot motions. The kinematics of orienting degrees of freedom has been chosen so that the tool pitch angles θ and the pitch directions γ are given separately by the degrees of freedom q_5 and q_4: $\theta(t) = q_5(t), \gamma(t) = q_4(t)$. In this case the laws of the generalized coordinates changes of the transportable degrees of freedom are described by the following:

$$q_1(t) = x(t) - l\sin(\theta(t))\cos(\gamma(t));$$
$$q_2(t) = y(t) - l\sin(\theta(t))\sin(\gamma(t));$$
$$q_3(t) = z(t) - l\cos(\theta(t)).$$

To simplify the controlling functions and increase the accuracy of tool positioning a special-purpose

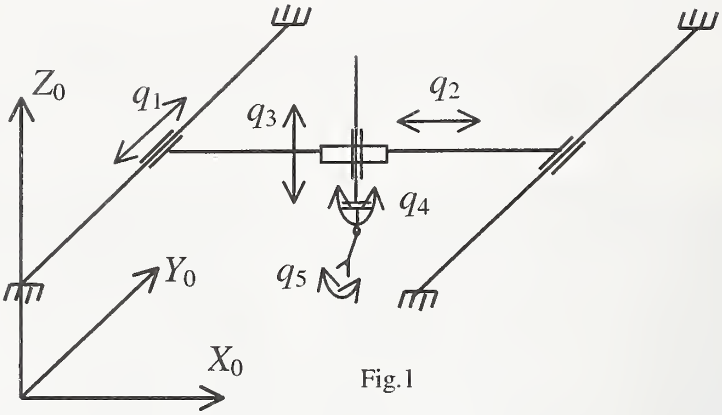


Fig.1

orienting head has been developed, its rotation axis of the last degree of freedom has been brought onto the plane of tool cutting. While rotating it allows to retain the center of the tool coordinate system on the rotation axis q_4 and obtain the main kinematics ratios of a simple kind:

$$[x(t), y(t), z(t), \theta(t), \gamma(t)] =$$
$$[q_1(t), q_2(t), q_3(t), q_4(t), q_5(t)]$$

Dynamic models have been used to control robot and develop control programs. For translation degrees of freedom in the basis of dynamic models the equation for the forces balance is used

$$\sum m_i \cdot \ddot{x} = F_e - \sum F_j ,$$

where m_i weights of movable parts; F_e - control force of degrees of freedom; F_j - disturbing effects incorporating frictional forces in movable parts of the mechanism and load forces on the working tool. Control force F_e is connected with a moment *M*. Of the drive by the ratio $F_e = M \cdot i_r \cdot G / r$, where i_r is the drive gear ratio; *G* is the gear box efficiency; *r* is an effective radius of the mechanism transforming rotational motions into translatory ones. The drive dynamics for each degree of freedom in models is presented in a linearized form by the system of equations

$$\alpha(s) = \frac{k_m}{s(T_m s + 1)} u(s),$$

where α is the angle of the motor shaft turn; *u* is the motor control voltage.

Degree of freedom along the coordinate *Z* has double function. One of them is connected with tool lifting and lowering and second - with creating necessary pressing force on the working surface. When resolved force control is applied to drive *Z* coordinate the relationships connecting the moment and control voltage have been incorporated in the model:

$$M_m(s) = \frac{k_m s(Js + f)}{s(T_m s + 1)} \cdot u_m(s),$$

where J, T_m, k_m are inertia, time constant and gain factor of the drive.

3. MOVEMENT PLANNING AND CONTROL

The specific feature of controlling robots for location and spacing of wood is the necessity of forming programmed trajectories of cutting tool movement. On their basis the prediction of displacements according to the coordinates is made and control voltages for each of them are detemined. In the foundation of the programmed robot control is the principle of setting movement trajectories for a cutting tool and movement program with help of the pattern being performed with CorelDraw vector graphic editor. While composing patterns their accurate scaling in a special file is provided, this file is then used as an assigning file while carrying out control. On the pattern for material location and spacing or for performing decorative cutting the coordinates of the initial point, and the starting point of the process trajectory are assigned, transition lines between closed figures of the pattern are assigned as well. While processing each figure pattern scanning with digitization step *T* and read-out of information about the coordinates of the next positioning $x[kT+1]$, $y[kT+1]$ are fulfilled. After the coordinates having been obtained control voltages are determined $u_x[kT+1]$ and $u_y[kT+1]$, and motion speeds along the coordinates *x* and *y* during the next control step are calculated as well. In order to reduce the effect of quantization as a control means in the interval $t_n \le t \le t_{n+1}$ we choose:

$$u[t_n] = 0.5[u(t_n) + u(t_{n+1})].$$

This ensures minimization of maximal deviation $\overline{u}[t_n]$ from the truth-value when monotonous change $u(t)$ takes place. When applying graphical means for setting movement trajectories there appears necessity to build algorithms ensuring the proper tool orientation in each point of the trajectory. To solve this problem different kinds of interpolation have been analysed and in the algorithms of robot control a parabolic interpolation in the interval $[(n-1)T,(n+1)T]$ is selected as a basic one. In this case for the time moment $t=(n+1)T$ a predictive calculation of coordinates for the point of tool position is made and according to the coordinates of three points coefficients of an interpolating equation are defined:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} x^2[(n-1)T] & x[(n-1)T] & 1 \\ x^2[nT] & x[nT] & 1 \\ x^2[(n+1)T] & x[(n+1)T] & 1 \end{bmatrix}^{-1} \begin{bmatrix} y[(n-1)T] \\ y[nT] \\ y[(n+1)T] \end{bmatrix}.$$

Next the obtained values of coefficients a_t are checked:

$$y[(n+1)T] = a_0 + a_1 x[(n+1)T] + a_2 x^2[(n+1)T]$$

and in case of identify the angle of normal to the trajectory in the point $t[(n+1)T]$ is calculated

$$\theta[(n+1)T] = 0.5\pi - \operatorname{arctg}(2a_0 x[(n+1)T] + a_1$$

After determining the vector of tool position and orienting and the vector of generalized coordinates $\overline{q}[(n+1)T]$ for the next step of control the values of tool motion speeds $\dot{x}$, $\dot{y}$ and the angular speed of rotation θ are calculated. The accuracy of development of the movement trajectory depends on the accuracy of fulfilling the task mentioned above. The simplest algorithm is the linear dependence

$$v_x[(n+1)T] = (x[(n+1)T] - x[nT])/T$$
$$v_y[(n+1)T] = (y[(n+1)T] - y[nT])/T$$
$$\omega_\theta[(n+1)T] = 2\pi(\theta[(n+1)T] - \theta[nT])/360 \cdot T$$

To perform high-quality patterns we have incorporated highly precision control algorithm. For this purpose the trajectory of tool movement is interpolated by a polynomial of the 3-d degree and according to the interpolating function $S_n(T)$ the derivative $\dot{S}_n[(n+1)T]$ is determined. The obtained derivative values are applied for calculating generalized speeds:

$$v_x[(n+1)T] = \dot{q}_1[(n+1)T] =$$
$$V_0 \cos(\operatorname{arctg}(\dot{S}_n[(n+1)T])),$$
$$v_y[(n+1)T] = \dot{q}_2[(n+1)T] =$$
$$V_0 \sin(\operatorname{arctg}(\dot{S}_n[(n+1)T])),$$
$$\omega_\theta[(n+1)T] = \dot{q}_4[(n+1)T] =$$
$$(\dot{S}_n[(n+1)T] - \dot{S}_n[nT])/T.$$

Taking into account the fact that during the functioning of control algorithm we can form sampled-data functions $q[nT]$ describing the trajectories of links movement, then to determine derivatives in the points of tool movement trajectory it is advantageous to use control algorithms for the model. While simulating the process of motion differentiation of digital sequences is presented as a sum of the kind:

$$\dot{q}[n] = T^{-1}\sum_{k=1}^{m} K^{-1}\nabla^k q[n] = \sum a_i q[n-i],$$

where $\nabla q[n] = q[n] - q[n-1]$ - inverse difference; m - number of terms of a degree series; $a_i = (-1)^i \sum_{i=0}^{m} K^{-1} C_k^i$; C_k^i - binomial coefficients.

Besides the considered control method on the graphical model it is necessary to include algorithms of programmed control by the path reference point into mathematical calculations and software. In this case reference points $P[x,y]$ and angles of tool orientation in each of them are assumed. Taking into account these values we form the data base to develop the laws of generalized coordinates changes. Motions planning thus is fulfilled on the basis of interpolation with cubic splines: $S_3(t) = \sum_{k=0}^{3} a_k t^k$.

Coefficients a_k are calculated in each section of interpolation having assumed that the trajectory is continuous and smooth. Planning of robot's movement is carried out with account of limitations in degrees of freedom which are preset with reference to the table of limit values for each coordinate $q_i^{\min} \le q_i \le q_i^{\max}$.

The authenticity of the measured is of great importance for control. In the encoding position sensors being used there may be abnormal short-Term imperfect data that can lead to short-term limit accelerations and deviations from the trajectory. Casual character of some charges of these sensors makes us to introduce predictive and correcting algorithms of simple calculating structure into the algorithms of robot control. In the applied algorithms expected values of position along the coordinates for the period of change T are predicted in each control step with help of m degree polynomial. Meanwhile we apply algorithms of single prediction, which is based on Lagrangian interpolating polynomial.

$$P(t) = \frac{1}{\tau^m} \sum P_{k-i} \prod_{j=0}^{m} \frac{k\tau + \tau r}{j-i} + \frac{M_{m+1}}{(m+1)!} \prod_{j=0}^{m} (k\tau + j\tau)$$

where k=1,2,3 – prediction step.

The analysis of prediction errors has shown that algorithms of double prediction should be applied for robots and values in three points (m=3) are to be taken. For convenience in usage of predictive expressions and reductions of calculations it is better to use recurrent form of the analysis.

A robot software solves the following problems: analog and discrete information about the parameters of robot condition is obtained, interface with drivers of analog inputs and outputs, output of discrete and analog control signals, formation of technological and emergency information, data display in real time. The programs can function in the environment of Windows and other versions and use all possibilities of this environment. Drivers for data exchange are formed as dynamic libraries DLL. The software includes robot models, which allow to carry out check of algorithm operation.

4. CONCLUSION

The material is prepared on the basis of the authors' research, which was carried out while developing robotic system for wood processing. The presented kinematic structure, algorithms of motions planning and control have been investigated on the models. Computer simulation of robot motions has shown the effectiveness of the described methods and algorithms. The obtained results of simulation were applied while developing and designing robots for cutting and spacing of wood.

SESSION 5

ECONOMIC ASSESSMENT OF AUTOMATION AND ROBOTICS TECHNOLOGIES

An Economic Assessment of Selected Integration and Automation Technologies

Robert E. Chapman

Office of Applied Economics, Building and Fire Research Laboratory (BFRL)
National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899
robert.chapman@nist.gov

Abstract: This paper presents a critical analysis of the economic impacts of past, ongoing, and planned research of BFRL's construction systems integration and automation technologies (CONSIAT) program. The CONSIAT program is an interdisciplinary research effort within BFRL to develop key enabling technologies, standard communication protocols, and advanced measurement technologies needed to facilitate the delivery of fully-integrated and automated project process (FIAPP) products and services to the construction industry. BFRL is participating in a public-private partnership focused on early commercialization of FIAPP products and services. The results of this analysis demonstrate that the use of FIAPP products and services will generate substantial cost savings to the capital facilities industry. The present value of savings nationwide expected from the use of FIAPP products and services in capital facilities over the next 15 years exceeds $3.5 billion. The present value of cost savings due to the public-private partnership's efforts focused on early commercialization of FIAPP products and services is expected to exceed $280 million.

Keywords: Building economics; commercial buildings; construction; impact evaluation; industrial facilities

1. INTRODUCTION:

The capital facilities industry[1] is facing several significant competitive challenges that will affect its future growth potential. Owners of capital facilities and contractors engaged in the construction of those facilities are pressing for reductions in delivery time as a means of improving their competitive positions. Owner concerns over both the first costs and the life-cycle costs of capital facilities and tightening profit margins for contractors are also affecting the competitive positions of each stakeholder. One means of improving the competitive position of each stakeholder in the capital facilities industry is through the development, adoption, and use of fully-integrated and automated project process (FIAPP) products and services.

FIAPP products and services offer the potential to reduce capital costs, reduce the delivery time of capital projects, and improve safety performance during construction. But investments in and the use of FIAPP products and services will be forthcoming only if the capital facilities industry perceives that the economic benefits outweigh the costs of using such products and services. Being able to demonstrate net economic savings from using FIAPP products and services will encourage their acceptance and use. The focus of this paper is on documenting how the use of FIAPP products and services in the capital facilities industry will generate significant net economic savings to the owners, operators, and managers of those facilities and to the contractors engaged in the construction of those facilities. These savings are based on results published in two economic impact assessments of BFRL's CONSIAT-related research and development effort [1, 2].

[1] The capital facilities industry covers construction-related activities and the associated supply chains throughout the life cycle of industrial facilities and commercial buildings. Industrial facilities include utilities, government facilities, and facilities where the manufacturing of products or commodities takes place. Commercial buildings include private- and public-sector office buildings, institutional buildings, and service businesses.

2. A CASE STUDY OF THE CAPITAL FACILITIES INDUSTRY:

The case study is divided into two sections. The first section focuses on data and assumptions. The second analyzes the cost savings from FIAPP products and services in the capital facilities industry.

2.1 Data and Assumptions

The base year establishes the anchor point for all calculations. The base year for computing all FIAPP-related costs and savings is 1997, a year for which authoritative and comprehensive construction industry cost data are available.

The diffusion of FIAPP products and services into the capital facilities industry employs two sets of diffusion models: one set for industrial facilities and one set for commercial buildings. Each set employs a primary diffusion model. The primary diffusion model, $P_\eta(t)$, gives the proportion of potential users who employ FIAPP products and services in time period t, where t = 1 corresponds to 2005, the anticipated time of first commercial use. In each set of diffusion models, the subscript η designates the market saturation level. The diffusion of FIAPP products and services into the marketplace is modeled up through 2017. By 2017, the use of FIAPP products and services is expected to be widespread. For more about the diffusion models, see [1, 2].

In order to estimate costs and savings due to the use of FIAPP products and services, it is necessary to specify both a base case and a FIAPP alternative. The term base case is used to represent the configuration that maintains the status quo (i.e., the "average" use of traditional design, information, and construction technologies). The FIAPP alternative is the configuration that provides equivalent or enhanced performance for all features of the base case through the use of FIAPP products and services.

There are two key differences between the two configurations. First, the degree to which construction activities (e.g., materials management) and facility service features (e.g., maintenance and repair procedures) are integrated, automated, and controlled is significantly higher in the FIAPP alternative. The second difference is that the FIAPP alternative has the potential to achieve enhanced performance for selected construction activities (e.g., better control of project cost and schedule) and facility service features (e.g., reduced maintenance and repair costs). These differences, although interrelated, are crucial in structuring differences in costs (e.g., due to the installation of additional equipment and software to generate improved systems integration, automation, and control) and savings (e.g., maintenance and repair cost savings due to the availability of electronic "as-built" information) between the two configurations.

The enhanced performance of the FIAPP alternative *vis-à-vis* the base case produces five types of cost savings. These cost savings are: (1) lower first costs; (2) lower maintenance and repair costs; (3) fewer construction-related accidents; (4) reductions in delivery time; and (5) higher net income for contractors. Lower first costs are registered through a reduction in a typical project's total installed costs (i.e., all project-related costs with the exception of land costs). Lower maintenance and repair costs are registered through reductions in future costs. Fewer construction-related accidents are registered through reductions in direct jobsite costs (e.g., medical costs), indirect jobsite costs (e.g., lost productivity of the crew due to the accident), and liability costs (e.g., claims costs). Reductions in delivery time are registered through increased opportunities for product sales and rental income. Higher net income for contractors is registered through the contractor's increased capability to control cost growth during the project delivery process. Cost savings accrue to owners and contractors in different ways. Lower first costs, lower maintenance and repair costs, and savings stemming from the earlier start-up of operations accrue to the owners and operators of capital facilities. Cost savings due to fewer construction-related accidents and higher net income accrue to contractors.

If capital facility owners, operators, and contractors employ the FIAPP alternative rather than the base case, they can expect to bear three types of additional costs. These costs are: (1)

higher evaluation costs; (2) increased costs of adapting new building products and services to industry use; and (3) increased training costs. These three costs may be classified as new-technology introduction costs. Ehlen and Marshall [3] define new-technology introduction costs as those costs covering the activities that bring the material/product from the research laboratory to full field implementation. New-technology introduction costs include the extra time and labor to design, test, monitor, and use the new technology. Ehlen's and Marshall's research on new-technology introduction costs is particularly relevant for this case study because they demonstrate that new-technology introduction costs disappear once the designer is satisfied with the technology's performance, the technology enters full implementation, and its application has become routine.

In performing the calculations presented in this case study, as well as those presented in [1, 2], a conservative approach to the estimation of savings and costs was employed. Potential cost savings are estimated based on project data from the Construction Industry Institute (CII) Benchmarking and Metrics Database.[2] These data were provided to NIST by CII as part of a research collaboration on the use of design/information technologies [4]. Because these technologies are commercially available now, they do not include the full potential for savings expected from FIAPP products and services. Thus, the cost savings reported here are lower-bound estimates of the savings expected from the use of FIAPP products and services. On the other hand, new-technology introduction costs are held constant throughout the study period. Since these costs are expected to decline over time, the estimated values for net economic savings reported here are also lower-bound estimates.

2.2 Analysis Results

Three types of information were combined to generate an estimate of cost savings nationwide. These three types of information are related to: (1) the diffusion models; (2) the cost savings due to reductions in first costs, maintenance and repair costs, and construction-related accidents *and* the increases in net income for owners and contractors due to reductions in delivery time and higher contractor profit margins; and (3) new-technology introduction costs. Estimates are produced for each year from 2005 to 2017. Each year's net cost savings was then discounted to a present value and summed to get the present value of cost savings nationwide. All present value calculations are based on standardized practices [5].

Table 1 summarizes how cost savings by category and in total are calculated. The years for which cost savings are calculated are listed in Column 1 of Table 1. Annual values for each category of cost savings are recorded in Column 2 for first costs, Column 3 for maintenance and repair costs, Column 4 for reductions in delivery time, Column 5 for higher net income for contractors, and Column 6 for construction-related accidents avoided. Note that no cost savings for any category occur until 2005, the year in which FIAPP products and services are expected to first become commercially available.

Reference to Columns 2 through 6 of Table 1 reveals different rates of change for cost savings. Cost savings increase slowly at first and then increase rapidly during the middle years (e.g., 2009 through 2013). The middle years of the study period correspond to the greatest rate of penetration of FIAPP products and services into the market place. Market penetration is modeled through application of two sets of diffusion models. As the rate of penetration into the market place slows, cost savings level off. In the case of the maintenance and repair category (see Column 3), cost savings peak in 2014 and then decline.[3] The differing rates of change have implications for net cost savings nationwide, which are presented in Table 2.

In addition to annual cost savings by category, Table 1 also contains total cost savings by year.

[2] All data provided to NIST by CII have been aggregated in a manner that precludes identification of an individual company's or project's performance.

[3] Reductions in maintenance and repair costs are measured on an annually recurring basis from the date of installation until the end of the study period in 2017. Thus, installations in the early (e.g., 2005) and middle years (e.g., 2012) have more years to generate savings than those occurring at the end of the study period.

These cost savings are recorded in Column 7. Total cost savings for each year equal the sum of each category's cost savings for that year. Total cost savings, denominated in millions of 1997 dollars, increase steadily between 2005 and 2016.

Table 2 summarizes how the present values of net cost savings nationwide by year and in total are calculated. The table also includes information on total cost savings, additional FIAPP-related installation costs, net cost savings, and the discount factor needed to translate yearly net cost savings into yearly present value cost savings nationwide. The years for which present values are calculated are listed in Column 1 of Table 2. Column 2 contains total cost savings by year in millions of 1997 dollars. The total cost savings for each year is transferred from the respective row of Column 7 of Table 1. The new-technology introduction costs associated with investments in FIAPP products and services for each year are recorded in Column 3 of Table 2. The difference between total cost savings and new-technology introduction costs equals net cost savings. Column 4 records net cost savings for each year in millions of 1997 dollars. Note that net cost savings increase steadily until 2015. The calculated value of the single present value factor for each year is recorded in Column 5. All entries are calculated using a real discount rate of 7 %. Because 1997 is the base year, the single present value factor takes on a value of 1.0 for that year. For years following 1997, the single present value factor is less than 1.0. The present value of net cost savings nationwide by year is recorded in Column 6. It equals the product of the net cost savings, in Column 4, and the single present value factor, in Column 5, for that year. Note that the present value of net cost savings nationwide increases steadily until 2015.

Because the entries in Column 6 are in present value terms, they can be summed to get total cost savings nationwide over the entire study period. Total cost savings nationwide resulting from the three sets of baseline analysis calculations are more than $3.5 billion ($3 532 million in present value 1997 dollars); see the bottom of Column 6 in Table 2.

Reference to Table 2 demonstrates the magnitude of the savings to the nation from using FIAPP products and services in the capital facilities industry. These cost savings nationwide also provide a basis for measuring the value of the public-private partnership's contribution. BFRL's dual role as a facilitator and developer of key FIAPP enabling technologies is expected to speed up the introduction of FIAPP products and services into the commercial marketplace. Because of the public-private partnership, FIAPP products and services are expected to be commercially available in 2005. Without a viable public-private partnership, the commercial introduction of FIAPP products and services is expected to be delayed until 2009. Information from subject matter experts and similar economic impact assessments suggest a range of values from two to five years for the likely delay. See [1, 2] to examine how variations in the likely delay were modeled.

Because the public-private partnership's efforts are expected to result in faster introduction of FIAPP products and services, those savings which would have been foregone in the event of a delay are attributable to the public-private partnership. Therefore, any savings over the first four years (starting with 2005), prior to the "delayed" introduction of FIAPP products and services in 2009, would have been foregone. Such an accounting framework may be handled through use of a 0 or 1 weighting factor. For those years in which savings are attributable to the public-private partnership, the weighting factor takes on a value of 1. The present value of those four-year's worth of savings exceeds $280 million, a strong indication of the value added of the collaborative efforts of the public-private partnership.[4]

3. CONCLUSIONS:

The $3.5 billion magnitude of national cost savings is impressive. Does it indicate, however, that investment in FIAPP products and services by individual owners and contractors will be cost effective? The answer to that question is almost certainly yes. Consider the case of the earliest

---

[4] CONSIAT-related research and development costs by BFRL are not included in the results presented in this article. Readers interested in how these costs were used to estimate the return on the public sector's CONSIAT-related investment are referred to [1, 2].

adopters of FIAPP products and services, those owners and contractors expected to invest in 2005. The aggregate investments, as measured by new-technology introduction costs, made by these owners/contractors is estimated as $24.0 million (see column 3 of Table 2). Their total cost savings are estimated as $71.6 million (see column 2 of Table 2). Thus, every dollar invested in 2005 generates nearly $3.00 in return. However, this is an incomplete picture, since the savings and costs accruing to owners and contractors are not evenly distributed. A more complete picture is provided through reference to Table 1. In Table 1, the different streams of cost savings are recorded according to whom they accrue. The owners of capital facilities capture the cost savings recorded in Columns 2, 3, and 4, whereas contractors capture the cost savings recorded in Columns 5 and 6. For early owner adopters, the entries in Columns 2 and 4 represent immediate, first-year cost savings of approximately $40.0 million. Thus, the entire first year's investment (i.e., new-technology introduction costs) can be covered by first-year cost savings captured by owners. Future owner cost savings due to reduced maintenance and repair expenditures add another $15.5 million. Contractor cost savings are lower than owner cost savings, but part of these lower savings are due to the more stringent values of incidence rates associated with the CII safety data versus industry incidence rates in calculating improved safety performance. Thus, it is likely that even if contractors bear a larger share of new-technology introduction costs, they will find investment in FIAPP products and services to be highly cost effective. These savings, coupled with the likelihood that new-technology introduction costs will decline over time, indicate that FIAPP products and services are an emerging technology whose time has come.

4. REFERENCES:

1. Chapman, Robert E. 2000. *Benefits and Costs of Research: A Case Study of Construction Systems Integration and Automation Technologies in Industrial Facilities*. NISTIR 6501. Gaithersburg, MD: National Institute of Standards and Technology.

2. Chapman, Robert E. 2001. *Benefits and Costs of Research: A Case Study of Construction Systems Integration and Automation Technologies in Commercial Buildings*. NISTIR 6763. Gaithersburg, MD: National Institute of Standards and Technology.

3. Ehlen, Mark A., and Harold E. Marshall. 1996. *The Economics of New-Technology Materials: A Case Study of FRP Bridge Decking*. NISTIR 5864. Gaithersburg, MD: National Institute of Standards and Technology.

4. Thomas, Stephen R., Candace L. Macken, and Sang-Hoon Lee. 2001. *Impacts of Design/Information Technology on Building and Industrial Projects*. NIST GCR 01-828. Gaithersburg, MD: National Institute of Standards and Technology.

5. American Society for Testing and Materials (ASTM). Fourth Edition, 1999. *ASTM Standards on Building Economics*. West Conshohocken, PA: ASTM.

Table 1. Cost Savings by Category and in Total by Year: 2005-2017

Year	Annual Cost Savings in Millions of 1997 Dollars by Category Due to					Total Cost Savings by Year in Millions of 1997 Dollars
	Reductions in First Cost	Reduced Maintenance and Repair	Reductions in Delivery Time	Higher Contractor Net Income	Improved Safety Performance	
Col. (1)	Col. (2)	Col. (3)	Col. (4)	Col. (5)	Col. (6)	Col. (7) (2)+(3)+(4)+(5)+(6)
2005	20.6	15.5	19.7	15.1	0.7	71.6
2006	49.9	29.6	47.2	36.7	1.6	165.0
2007	79.9	40.5	74.5	58.7	2.4	256.0
2008	125.5	62.7	115.6	92.2	3.6	399.7
2009	192.1	94.5	174.7	141.1	5.2	607.6
2010	283.2	137.1	254.3	208.1	7.3	890.1
2011	397.6	189.6	352.6	292.1	9.7	1 241.7
2012	526.5	247.1	461.3	386.8	12.2	1 634.0
2013	655.5	299.3	567.4	481.6	14.4	2 018.2
2014	769.9	330.9	658.5	565.6	16.0	2 340.8
2015	861.0	323.1	727.7	632.5	17.0	2 561.3
2016	927.5	261.3	774.9	681.4	17.3	2 662.5
2017	973.2	146.8	803.6	715.0	17.2	2 655.8

Table 2. Present Value Cost Savings Nationwide by Year and in Total

Year	Total Cost Savings in Millions	New-Technology Introduction Costs	Net Cost Savings in Millions by Year	Single Present Value Factor by Year	Present Value of Net Cost Savings Nationwide by Year in Millions
Col. (1)	Col. (2)	Col. (3)	Col. (4) (2) - (3)	Col. (5)	Col. (6) (4) x (5)
2005	71.6	24.0	47.6	0.582	27.7
2006	165.0	58.3	106.7	0.544	58.0
2007	256.0	93.3	162.7	0.508	82.7
2008	399.7	146.6	253.1	0.475	120.2
2009	607.6	224.4	383.3	0.444	170.2
2010	890.1	330.8	559.3	0.415	232.1
2011	1 241.7	464.4	777.3	0.388	301.5
2012	1 634.0	615.0	1 019.0	0.362	369.3
2013	2 018.2	765.6	1 252.6	0.339	424.3
2014	2 340.8	899.2	1 441.7	0.317	456.4
2015	2 561.3	1 005.6	1 555.7	0.296	460.3
2016	2 662.5	1 083.4	1 579.1	0.277	436.6
2017	2 655.8	1 136.7	1 519.1	0.258	392.6
				TOTAL	3 531.9

A Micro Level Analysis Of The Relationship Between Changes In Equipment Technology And Wages In The U.S. Construction Industry

by

Paul M. Goodrum [1] and Manish Gangwar [2]

ABSTRACT:

A shortage of skilled labor exists in the construction industry. Fortunately, advancement in construction equipment and material technologies, along with modularized components and estimating and scheduling strategies have offset the shortage of skilled construction labor. The construction industry has witnessed a drop in real wages since 1970. Decline in real wages may be attributed to a combination of socioeconomic factors like migrant laborers, fringe benefits, safety procedures, union membership and worker skills. Another factor that may be impacting construction real wages is technological changes over the past couple of decades; including technological changes in construction equipment. There is a growing need to understand how changes in technology are affecting employment conditions in construction. If more could be known about how technology affects wages, the industry could formulate better strategies for future workforce needs. This paper examines the relationship between changes in equipment technology and changes in construction wages with the help of five factors of equipment technology change; control, energy, ergonomics, functionality and information processing. Furthermore, data from the U.S. Bureau of Labor Statistics' Current Population Survey (CPS) is used to examine the effects of computer usage on wages among hourly workers in construction.

KEYWORDS:

1. INTRODUCTION:

The U.S. construction industry contributes significantly to the U.S. economy. When one includes construction related business involving design, equipment and materials manufacturing, and supply, the construction industry accounts for 13% of the GDP, making it the largest manufacturing industry in the U.S. (BEA 2000).

The shortage of skilled workers is considered to be one of the greatest challenges facing the U.S. construction industry. Not since the early 1970s and post World War II has the U.S. construction industry experienced such low unemployment rates (BLS 2002). Advances in construction equipment and material technologies, modularized components, and estimating and scheduling strategies have offset the shortage of skilled construction labor. However, there is a perception among industry leaders that the skilled worker shortage is getting worse. A survey of facility owners showed that 78% thought the skilled worker shortage had increased during the past 3 years (Rosenbaum 2001).

Although real wages in general in the U.S. began to outpace inflation in the late 1990's, there has been a long-term decline in construction real wages since the 1970's (Allmon, et al. 2000 and Oppedahl 2000). Other industries, such as manufacturing, have also experienced declines in real wages; however, the declines have typically been greater in construction. This greater decline may be due to a combination of socioeconomic factors including an increase in migrant laborers in construction, fringe benefits, and construction safety, and a decrease in union membership and worker skills (Oppedahl 2000, Goodrum 2002).

Another factor that may be impacting construction real wages is technology. Over the past couple of decades, there has been a wide array of technological changes in construction equipment and material technology. Construction equipment has become more powerful, automated, more precise, safer, and more functional, allowing workers to be more productive in construction activities. In many instances, technology has made construction equipment easier to use. One example is heavy machinery. Advancements in hydraulic controls

1. Assistant Professor, Department of Civil Engineering, 151C Raymond Bldg., University of Kentucky, Lexington, KY 40506-0281 email: pgoodrum@engr.uky.edu

2. Graduate Research Assistant, Department of Civil Engineering, 151D Raymond Bldg., University of Kentucky, Lexington, KY 40506-0281

and microprocessors have automated and simplified the operation of earthmoving machinery. There have also been advancements in construction equipment that have introduced new technologies that require skill sets normally outside those traditionally required for construction. For example, the use of Global Positioning Systems onboard earthmoving equipment now require equipment operators to be proficient in the use of computers.

2. METHODOLOGY

This paper examines the effect of equipment technology on construction wages in two parts. First, the effects of changes in equipment technology on real wages from 1976 to 1998 are examined. This involves examining the changes in five technology factors (Amplification of Human Energy, Level of Control, Functional Range, Ergonomics, and Information Processing) and the change in the average wage of workers in crews for 100 construction activities. Second, the effects of computer usage on construction wages are examined for 470 individual hourly construction workers.

2.1. Equipment Technology Defined

This research examines the effect of changes in equipment technology on construction wages, specifically the equipment technologies of hand tools, machinery, and computers. Hand tools include pneumatic nail guns, electric drills, circular saws, and similar types of tools. Machinery includes cranes, grout pumps, bulldozers, and similar types of implements.

2.2. Technology Factors

To examine how different mechanisms of equipment technology change have influenced construction wages, five factors were identified (defined below and examples discussed later) to characterize changes in technology.

Amplification of Human Energy: technology designed to make an activity easier to perform physically. In its simplest terms, it can be regarded as the shift in energy from human to machine bringing an increase in energy output.

Level of Control: advances in machinery and hand tools that transfer control from human to machine.

Functional Range: changes that expand a tool or machine's range of capabilities.

Ergonomics: technology that alleviates physical stresses imposed on a worker and helps the worker cope with the work environment

Information Processing: over time, construction equipment has been designed to provide greater and more accurate information regarding internal and external processes. This factor includes the incorporation of computers into the work processes.

3. DATA SOURCES

3.1. Estimation Manual

The data for the research came from the estimation handbook Means Building Construction Cost Data (Means) and the Computer and Internet Use Supplement, data files for 2001 from the U.S. Bureau of Labor Statistics' Current Population Survey. Wage data from the 1976 and 1998 Means estimation handbooks on 100 activities was collected to examine the effects of changes in equipment technology (as defined by the technology factors) on construction wages. Data from the CPS was used specifically to examine the effects of the use of computers on construction wages.

These estimation handbooks provide wage data, unit labor costs, unit equipment costs, physical output data, and work-hour requirements for construction activities. While the handbooks are a valuable source of information about construction cost and productivity across time, there are some limitations to the data. The contractors who provide the figures for the manuals are not required to build a project using their estimations; this leads some contractors to submit inflated estimates of construction costs (Pieper 1989).

Three criteria were used to select activities for inclusion in the study. The first criterion was that the same activity be found in both the 1998 and 1976 estimation manuals. Due to changes in methodology, materials, or lack of use in construction, a number of activities included in the 1976 manual were not included in the 1998 manual. Likewise, a number of new activities were included in the 1998 manual due to new methodology or materials. Second, activities from a diverse range of technological changes were selected. Third, activities were selected to represent a wide range of activity types from different divisions of the Construction Specification Institute (CSI) master format.

3.2. CPS September 2001 Computer and Internet Use Supplement

To further examine the effects of computer usage on construction wages, data was collected from the September 2001 Computer and Internet Use Supplement from the U.S. Bureau of Labor Statistic's (BLS) Current Population Survey (CPS). The CPS is a monthly survey of approximately 50,000 households conducted by the U.S. Census Bureau for the U.S. Department of Labor. With the survey being conducted for more than 50 years, CPS data provides information on economic indicators, which influence U.S. governmental policy. Data from the CPS is available to the public via their website. (http://www.bls.census.gov/cps/cpsmain.htm).

Each month, the CPS randomly selects 59,000 housing units (e.g. single family homes, townhouses, condominiums, apartment units, and mobile homes) for the sample, and approximately 50,000 are occupied and eligible for the survey. The other units are found ineligible because they have been destroyed, vacant, converted to nonresidential use, or contain persons whose usual place of residence is elsewhere. Respondents are asked questions about the employment information and demographic characteristics of each member of the household over 14 years of age. In September 2001, the Computer and Internet usage survey was added as a supplement to that month's CPS. In addition to the demographic data collected each month, the Computer and Internet Supplement contained questions about the respondent's use of computers, including the use of computers at work, which was used in the research's analysis.

A number of criteria were used to select cases (each case representing an individual respondent) from the September 2001 CPS Computer Supplement data. First, only individuals listing their primary industry of employment as construction were selected. Next, each case had to meet the following series of additional selection criteria:

1. Full-time hourly workers;
2. Male construction workers;
3. Non-supervisory construction workers;
4. Hourly wage greater than or equal to the U.S. minimum wage of $5.15/hour.

The use of these selection criteria resulted in 470 cases.

4. ANALYSIS

4.1. Effects of Changes in Equipment Technology on Real Wages from 1976 to 1998

4.1.1. Measured Change in Equipment Technology

The authors identified and examined 43 types of hand tools and 31 types of machinery in the 100 construction activities. Obviously, many hand tools and machinery were used in several activities. Equipment technology changes were identified using equipment catalogs, handbooks and specifications. Figure 1 shows the number of activities that experienced a change in equipment technology in at least one tool or item of machinery for each of the technology factors.

As shown in Figure 1, more than 70% of the activities experienced an increase in energy output. Prior related research indicates that the metals, wood and plastic, and site-work divisions experienced the greatest amount of change in tool and machinery energy output (Goodrum and Haas 2002). One example of change in energy output in the metals division involves welding machines, which offer increased wattage output. The powder actuated systems in the metals divisions used in metal decking offer greater depth penetration for installed studs. In addition, by 1998 cranes offered more lifting capacity than available in 1976. In the wood and plastic division, circular saws operated at higher RPMs, and the pneumatic nail gun required less human energy than a hand held hammer. Most site work machinery increased in horsepower output including front-end loaders, dump trucks, backhoes, bulldozers, graders, asphalt pavers, and scrapers.

As seen in Figure 1, almost half of construction activities experienced a change in the amount of human control needed from 1976 to 1998. Welding machines in the metals division, for instance, are now equipped with remote controlled amperage adjusters and powder actuated systems have semi-automatic loading capability. The pneumatic nail gun has replaced the hand held hammer in the woods and plastic division and in formwork installation in the concrete division. Also in the concrete division, pump trucks are now equipped with remote controlled booms, and concrete vibrators automatically adjust the vibration frequency to match the concrete's slump.

Changes in functional range occurred in slightly less than half of the activities (Figure 1). Through advancements in hydraulic controls and microprocessors, site-work machinery now has greater precision and a longer reach for booms and buckets. Excavators and backhoes are capable of digging deeper.

Figure 1 shows that exactly half of the construction activities experienced some change in ergonomics. For example, by 1998 many hand tools, such as circular saws, hand drills, pneumatic nail guns, and caulking guns, were lighter and operated with less noise and vibration than their predecessors.

Almost all of the advances in information processing occurred in heavy machinery (Goodrum and Haas 2002). This finding explains why most construction activities did not experience such an improvement in equipment technology. For example, some heavy machinery now offer self-monitoring and self-diagnostic systems.

4.1.2. Measured Change in Real Wages

Daily crew wages as reported in Means were divided by the number of crewmembers in each activity to estimate individual worker's daily wage. In order to measure real wages (wages adjusted for inflation), the Census Construction Cost Index was used to normalize wages to 1990 levels. A description of the Census Construction Cost Index can be found at the Department of Commerce website (http://www.census.gov/prod/3/98pubs/c30-9805.pdf).

The overall average change from 1976 to 1998 in a worker's daily real wage was -$19.97, with a 95% confidence interval of ± $6.97. This confirms other findings that show a long-term decline in construction real wages (Allmon, et al. 2000, Oppedahl 2000). Figure 2 illustrates the average changes in daily real wages for each division of the CSI Master format.

On average, concrete activities experienced the largest decline in daily real wages, while masonry activities experienced little change. Further research is needed to determine the reasons behind the various sector changes.

4.1.3. Relation Between Equipment Technology and Partial Factor Productivity Change

Analysis of Variance (ANOVA) is used to test whether two or more groups have statistically significant different means. The ANOVA test estimates the statistical significance of the difference between the means (F-value), and it measures the amount of variation in the dependent variable that is explained by the independent variable Eta Square (e). The ANOVA analyses compared the daily real wage changes from 1976 to 1998 for (1) activities that experienced a change according to the technology factor and (2) activities that had not. Figure 3 shows the ANOVA results.

With the exception of energy and ergonomics, the activities that observed a change in equipment technology experienced a statistically significant different decline in daily real wages. Activities with an equipment change in functional range and information processing experienced over 60% less of a decline in daily real wages compared to activities without such changes. One possible explanation for these differences is the added skills required for workers to adopt these types of equipment technology changes, which may result in higher wages. Activities experiencing a change in level of control actually experienced over 150% more of a decline in real wages compared to activities without change. A possible explanation for this added decline is that many changes in level of control serve to simplify the processes, which may result in lower wages. Further research in the area is needed to examine other reasons.

4.2. Effects of Computer Usage on Construction Wages

One result of the previous set of analyses was that information processing has a substantial and significant relation with activities that saw less of a decline in daily real wages compared to activities that did not experience such a change. Because this phase of the study was limited to examining changes in equipment technology that were widely diffused in construction, most of the changes in information processing were found only in heavy machinery. To further examine how changes in information processing affect construction wages, data from the CPS September 2001 Computer Supplement was analyzed.

4.2.1 Measured Computer Usage Among Non-Supervisory Construction Workers

Of the 470 cases analyzed in the CPS September 2001 Computer Supplement, 49 (10.4%) indicated they used a computer at work. The top three occupations that used computers were: (1) electricians, (2) electrical power installer and repairers, and (3) plumbers. Occupations in which there were no respondents indicating they used computers as work included: roofers, concrete and terrazzo finishers, electrician apprentices, hard and soft tile setter`s, insulation workers and sheet metal duct installers. Unfortunately, the Computer Supplement data did not measure how the computers were used at work.

4.2.2. Relation Between Computer Usage and Wages in Construction

Data was analyzed from the CPS September 2001 Computer Supplement to examine the effects of computer usages on construction wages by comparing hourly wages between construction workers who use a computer at work and those who do not use computer at work (Figure 4). The difference in education, work experience, and age was also examined between those who do and do not use a computer at work.

Information from the CPS is used to create more than 350 variables. The CPS, however, does not ask respondents about their work experience, an important consideration in a study on wage differentials. One method for estimating work experience, used by the BLS, is to use CPS data to calculate potential experience using the following equation (1) (U.S. Department of Labor. (1993)). The units of potential experience are given in years.

$$\text{Potential Experience} = \text{Age} - 6 - \text{Years of School} \quad (1)$$

Variable for education was recoded by the researchers to represent number of years of education completed at school. Women's work experience is found to be substantially influenced by being married and having children. To avoid these influences, this study focused on men.

These analyses show that non-supervisory construction workers who use computers at work are significantly paid more than workers who do not use computers at work (the average hourly wage among workers who use computers was $18.43 compared to $15.56 for those who did not). At the same time, workers who use computers at work are statistically significantly more experienced (workers who used computers had on average 22 years of experience compared to 18 years of experience for those who did not); more educated (workers who used computers had on average 12.8 years of education compared to 11.6 for those who did not); and older (workers who used computers were on average 40.8 years old compared to 35.7 years old for those who did not). Although this analysis indicates a relation between higher wages and the use of computers for non-supervisory construction workers, it is not clear whether the increase in average hourly wage is due to usage of computer or merely a reflection of already established relations with experience, education and age.

5. CONCLUSIONS:

The findings reported here indicate that:

1. The decline in real wages exists throughout all sectors and divisons in construction.
2. Activities that experienced a change in Functional Range and Information Processing experienced less of a decline in real wages compared to activities that did not.
3. Not all changes in equipment technology are related to lessened declines in real wages. Activities that experienced a change in Level of Control actually experienced greater declines in real wages.
4. Non-supervisory construction workers who use computers at work earn higher hourly wages, although further research is needed to account for the effects of experience, education, and age.

6. REFERENCES:

1. Allmon, E., Haas, C., Borcherding, J., Goodrum, P. (2000). "US Construction Labor Productivity Trends, 1970-1998." J. Constr. Engrg. And Mgmt., ASCE. 126(2), 97-104
2. Bureau of Economic Analysis (BEA) Website (2000). http://www.bea.doc.gov Accessed April 2002
3. Bureau of Labor Statistics (BLS) 2002, Current Employment Statistics Survey. http://www.bls.gov/ces, accessed June 13, 2002.
4. Goodrum, P., and Haas, C. (2002). "Partial Factor Productivity and Equipment Technology Change at the Activity Level in the U.S. Construction Industry." J. Constr. Engrg. And Mgmt., ASCE. (Forthcoming).
5. Means Company, Inc (19XX). Building Construction Cost Data. Philip Waier, Sr. Editor, Kingston, MA.
6. Oppedahl, D. (2000). "Understanding the (relative) rise and fall of construction Real Wages." Chicago Fed Letter. (155): 1-4.2000 Jul
7. Pieper, P. (1989). "Construction Prices Statistics Revisted" in Jorgenson, D. and Landau, R. Technology and Capital Formation
8. Rosenblum, (2001). "Craft Labor Shortage Provokes More Study of Pay and Safety." Engineering News Read, McGraw Hill, August 20 2001. P11.
9. Slater, S., ed. (1997). Business statistics of the United States, Berna Press, Lanham, Md. U.S. Department of Labor. (1993). Labor Composition and U.S. Productivity Growth, 1948-90. U.S. Government Printing Office, Washington, D.C.

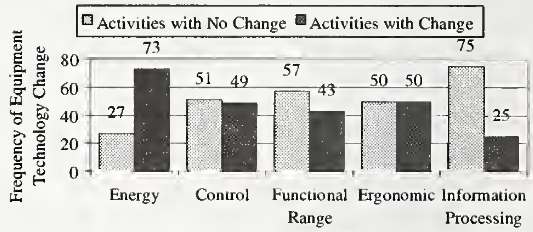


Figure 1: Change in Equipment Technology by Tecnology Factors, 1976 - 1998

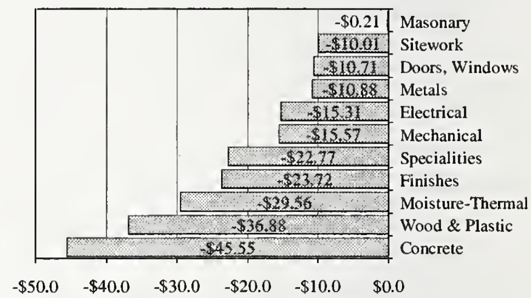


Figure 2: Change in Daily Real Wages (1990$) by Division (Data Source: Means 19XX)

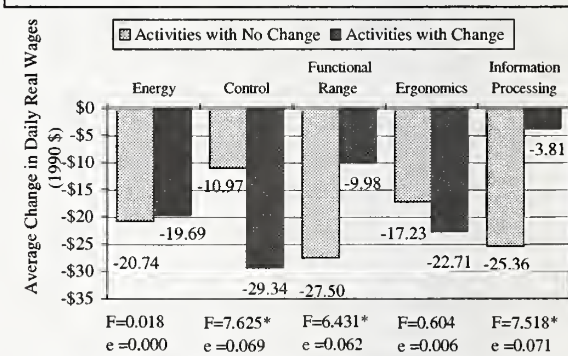


Figure 3: A Comparison of Daily Real Wages Change from1976 to 1998 for Activities that Experienced a Change in Equipment Technology and Activities with no Change. *p > 0.05

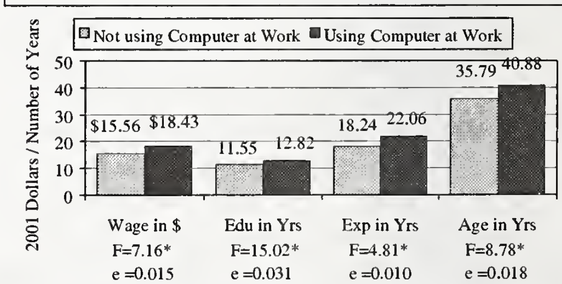


Figure 4: Difference in Average Wages, Average Education, Average Experience and Average Age Between Non-Supervisory Construction Workers Who Use Computer at Work and Workers Who Do Not. *p > 0.05 (Data Source: CPS 2001 Supplement for Computer and Internet Usage)

Generating and Maintaining Activity-based Cost Estimates with Feature-Based Product Models

by

Sheryl Staub-French[1] and Martin Fischer[2]

ABSTRACT

Understanding how the building design influences construction costs is a challenging task for estimators. Estimators must recognize the design conditions that affect construction costs and customize the cost estimate accordingly. Estimators have different preferences for how and when to adjust a project's activities, resources, and resource productivity rates that form the basis of a cost estimate. Current tools and methodologies lack ways to help estimators customize construction cost information according to their preferences and maintain cost estimates as the design changes based on those preferences. This paper describes the activity-based cost estimating process we formalized to help estimators customize a project's activities, resources, and resource productivity rates based on their preferences and the particular features in a given product model. We implemented and tested the process in a prototype called Activity-based Cost Estimating (ACE). ACE creates a set of project-specific activities that know why they are needed in the cost estimate, what feature requires their execution, what resources are executing the activity and why, and what their labor and material costs are. Our tests show that ACE helps estimators to generate and maintain cost estimates more completely, consistently, and quickly than state-of-the-art cost estimating software.

KEYWORDS: Cost Estimating; Information Technology; Construction; Product Features

1. INTRODUCTION

It is the cost estimator's task to determine how a building design influences construction costs. Estimators must determine what design conditions are important (i.e., incur a cost), when they are important, and how they affect construction costs when creating cost estimates. Construction cost estimates are used to assist designers, owners, and builders of facilities in resolving a variety of decisions, such as evaluating the cost of different design alternatives, budgeting construction costs, and establishing the cost impact of design changes. Consequently, it is critical that cost estimators provide detailed and accurate cost estimates in a timely manner to support project teams in making these different decisions.

Estimators using state-of-the-art estimating software can establish a relationship between a component in a product model and a cost item in a cost-estimating database when creating a cost estimate (Timberline 2001). These relationships help estimators to take off quantities automatically by representing the component properties that affect construction costs. However, this representation is incomplete because it does not represent the estimator's rationale for how the component properties affect specific cost information, and it does not represent other design conditions and their impact on a component's cost. Consequently, generating and maintaining cost estimates today is a largely manual, error-prone, and time-consuming process.

This paper describes the formalization of an activity-based cost estimating process that leverages the rich representation of standard product models to help estimators generate and maintain construction cost estimates. This research is based on a completed research project and was motivated by our experiences with state-of-the-art integrated estimating software on real projects. This work addresses

[1] Assistant Professor, Department of Civil Engineering, University of British Columbia, 2324 Main Mall, Vancouver, BC, Canada, V6T 1Z4. Voice: 604-827-5118, Fax: 604-822-6901, sherylsf@civil.ubc.ca.

[2] Associate Professor, Department of Civil and Environmental Engineering and (by Courtesy) Computer Science, Director, Center for Integrated Facility Engineering, Terman Engineering Center, Room 292, Stanford University, Stanford, CA 94305-4020. Voice: 650-725-4649, Fax: 650-725-6014, fischer@stanford.edu.

conference themes related to computer integrated building processes and construction process modeling and simulation.

2. CASE STUDY

This section describes a case study to illustrate the requirements for automated support of the cost estimating process. The case study is based on a drywall estimator's process for estimating the labor costs for one of the rooms in an office project shown in Figure 1. Figure 2 shows an estimator's rationale for adjusting the activities, resources, and resource productivity rates to account for the cost impacts of specific design conditions.

The case study demonstrates that there are different types of design conditions that affect construction costs. Design conditions can be based on properties of components (e.g., the 'curvature' of the wall), intersections of components (e.g., the 'structural penetration' resulting from the intersection of the wall and beam), and groupings of components (e.g., the 'grouping of walls' based on component similarity). The case study also demonstrates that design conditions can affect construction costs in different ways. Design conditions can affect the requirement for activities (e.g., the 'structural penetration' requires the activity Apply Caulk), design conditions can affect when a resource is appropriate in an activity (e.g., the height of the wall affects the need for Rolling Scaffolding), and design conditions can affect a resource's ability to execute an activity effectively (e.g., the similarity of the walls leads to an increase in crew productivity).

Estimators using state-of-the-art cost estimating software cannot represent many of the design conditions that affect construction costs (e.g., structural penetrations and component similarity). Furthermore, estimators using state-of-the-art cost estimating software cannot explicitly represent the specific cost information affected and the way it is affected by different design conditions (e.g., resource use or execution). Consequently, estimators have to manually identify most design conditions and manually adjust the project's activities and resources accordingly. For a large project, it is typically too time-consuming to make these project-specific adjustments manually for all the different design conditions in a given product model. Consequently, estimators often employ ad hoc methods (e.g., adjusting the crew productivity rates of all the "Install Metal Studs" activities to account for one wall's curvature) and overlook the cost impact of different design conditions (e.g., overlook cost impacts resulting from openings). Moreover, estimators often do not have the time to provide specific feedback to designers on the cost implications of their design decisions (e.g., the cost implications of a specific wall height). The lack of a formal process and automated support leads to inconsistencies and inefficiencies in the cost estimating process and resulting cost estimate, and limits the ability of estimators to help designers develop cost-effective designs.

3. GENERATING AND MAINTAINING COST ESTIMATES USING ACE

The activity-based cost estimating process we formalized helps estimators to generate and maintain cost estimates quickly and consistently based on the design conditions in a given product model. We implemented and tested this process in a prototype called Activity-based Cost Estimating (ACE). The main challenges associated with providing automated support of the cost estimating process are that different design conditions exist in any given product model, that different design conditions affect construction costs in different ways, and that estimators have different preferences for how and when to adjust construction costs to account for different design conditions.

We developed the activity-based cost estimating process by abstracting the design conditions estimators consider and the different ways estimators adjust activities and resources to account for different design conditions. We use *features* to describe the specific part of the design that estimators care about, *design conditions* to describe when features are important to estimators, and *activities* to describe how features impact production and hence, construction cost. By leveraging the activity-based cost estimating process, ACE allows estimators to represent their rationale for relating features of a building product model with construction activities to enable automated support of the cost estimating process.

Figure 3 shows the three different modules of the activity-based cost estimating process:
(1) Instantiate Features: Identify the relevant cost-driving features in the input product model and instantiate the features to create an estimator-focused feature-based product model.
(2) Customize Activities and Resources: Customize the activities and resources for each component being estimated based on the estimator's rationale and the particular features in the estimator-focused feature-based product model.
(3) Generate and Maintain Construction Costs: Calculate each activity's quantities and duration to determine the activity's cost. If the estimate is based on a revised design, identify the cost information affected and reconcile the activities and resources so that the design and estimate remain in balance. The output of this process is a set of resource-loaded and cost-loaded activities that are explicitly related to the features in the input product model.

3.1 M1: Instantiate Features

The motivating case shows that different types of design conditions affect construction costs. Estimators consider properties of components, intersections of components, and groupings of components when creating cost estimates. The purpose of the first module is to transform designer-focused product models into feature-based product models that support cost estimating. The input product model is represented using the industry standard Industry Foundation Classes (IFC) (IAI 2001). ACE identifies the cost-driving features in the input IFC-based product model to create an estimator-focused feature-based product model.

The IFC's explicitly represent components, attributes of components, and relationships between components in building product models (IAI 2001). However, they do not explicitly represent many of the design conditions that affect construction costs, such as penetrations and component similarity. We use features to represent the design conditions that are important to cost estimators of building construction. Product features are used extensively in manufacturing to describe the geometric forms or entities in a product model that are important in some aspect of the manufacturing process (Cunningham and Dixon 1988). However, the feature representations developed in the manufacturing industry do not fully support the representation of building product models. Specifically, building product models contain different features and different types of products. Our research applies the manufacturing concept of features to building construction and extends it to represent the features that are useful to cost estimators.

We modeled three different types of features in this research: (1) component features (e.g., 'walls'), (2) intersection features (e.g., 'structural penetrations'), and (3) macro features (e.g., 'groupings of components based on similarity'). We formalized a feature ontology that represents the different attributes of each feature type and enables estimators to represent feature instances according to their preferences (Staub-French 2002). The feature ontology provides the map to relate an IFC-based product model to an estimator-focused product model. We represent features in a project-independent way so that they can be reused from project to project to identify the relevant features given an IFC-based product model.

ACE leverages the feature ontology to provide a framework for estimators to represent their preferences for naming features, specifying relevant component intersections, defining component similarity, and specifying the features that affect a specific component's construction costs. For example, the estimator from the motivating case can represent the "structural penetration" as a feature that results from the intersection of 'walls' and 'beams,' and specify that this feature is important for constructing 'walls.' ACE analyzes the geometry and topological relationships between the components in the input IFC-based product model to identify the cost-driving features specified by the estimator. Hence, ACE enriches current standard building product models by representing the features of building product models that affect construction costs.

3.2 M2: Customize Activities and Resources

The case study shows that estimators adjust the project's activities, resources, and resource productivity rates to account for the cost impact of different features. The purpose of the second module is to help estimators customize the activities and resources in a cost estimate according to their preferences based on the specific features in a given product model.

Prior research efforts demonstrate that cost estimates can be generated directly from 3D models (Laitinen 1998; Aouad et al. 1994; Aouad et al. 1997, Staub-French and Fischer 2001). However, these research efforts do not customize construction cost information for specific design conditions in a given product model, and they do not account for different estimator preferences. Other research efforts recognize design conditions that affect construction costs and customize the activity's resources and resource productivity rates accordingly (Fischer 1991; Thomas and Zavrski 2000; Hanna et al. 1992). However, these research efforts do not represent and account for different estimator preferences when customizing the project's resources and resource productivity rates to the design conditions in a particular product model.

We abstracted the common attributes of estimators' rationale for how and when different design conditions affect construction costs and developed templates to capture this estimating knowledge from estimators (Staub-French 2002). The templates allow estimators to specify the features that affect activities (Activity Specification templates) and the features that affect resources (Resource Specification templates). The templates provide a structured way for estimators to represent the specific impact different features have on a project's activities and resources. Estimators input their rationale once in the Activity and Resource Specification templates and ACE reuses this knowledge from project to project when generating and maintaining cost estimates.

In ACE, we implemented a formal process that automatically customizes activities and resources when generating and maintaining cost estimates for estimator-focused feature-based product models (Staub-French 2002). For each feature in the input product model, ACE identifies the relevant Activity Specifications and adds the specified activity to the estimate. Then, ACE assigns resources to the activities and adjusts the resources' productivity rates according to the estimator's preferences in Resource Specifications and based on the specific features in the feature-based product model. The output of the second module is a set of project-specific resource-loaded activities that are explicitly related to the estimator-focused feature-based product model and the estimator's rationale.

3.3 M3: Generate and Maintain Construction Costs

The purpose of the third module is to generate and maintain construction costs given the input resource-loaded activities and related features. Calculating the construction costs for resource-loaded activities is a straightforward process. However, the explicit relationships between features, activities, resources, costs, and the estimator's rationale in Activity and Resource Specifications enable the maintenance of cost estimates if the design changes.

Many research efforts have developed computer tools that automatically generate the project-specific relationships between components, activities, resources, and costs (Laitinen 1998; Aouad et al. 1994; Aouad et al. 1997; Froese 1992). However, they do not represent why components, activities, resources, and costs are related and when the relationships are needed.

Our research extends existing formalisms of construction processes that define activities as objects <O>, actions <A>, and resources <R> (Darwiche et al. 1988; Aalami 1998). Our research extends this formalism by generating activities that also know what feature <F> requires the activity's execution and how much the activity costs <C> to create an integrated <FOARC> model consisting of activities that explicitly relate features, objects, actions, resources, and costs.

Each activity generated by ACE knows what feature requires its execution, the material and resource cost implications of the activity, the estimator's rationale for adding the activity, what resources are executing the activity and why, and how particular features affect the resources' productivity rate. Consequently, ACE can help estimators to identify the cost information affected by design changes and calculate the corresponding cost impact of design changes. Our tests show that the cost estimating process implemented in ACE and the resulting integrated <FOARC> model enables estimators to generate and maintain construction cost estimates from feature-based product models more completely (i.e., less ad hoc and with fewer omissions), consistently, and quickly (Staub-French 2002).

4. CONCLUSIONS

This paper described an activity-based cost estimating process that creates an integrated model consisting of activities that explicitly relate features, objects, actions, resources, and costs, and the estimator's rationale for relating this information. This process helps estimators to generate and maintain construction cost estimates and avoid ad hoc and error-prone methods that lead to inconsistencies and inefficiencies in the cost estimating process.

The formalisms developed and implemented in ACE take an essential step toward creating software tools that can help project teams to maintain integrated models of a project's scope, schedule, and cost. Understanding the relationships between this information is critical to managing the design and construction process. Future research directions should address other types of features and factors exogenous to product design, such as site characteristics and resource skill and availability.

5. REFERENCES

Aalami, F. (1998). "Using Method Models to Generate 4D Production Models," PhD Thesis, Stanford University, Stanford.

Aouad, G., Betts, M., Brandon, P., Brown, F., Child, T., Cooper, G., Ford, S., Kirkham, J., Oxma, R., Sarshar, M., and Young, B. (1994). "ICON: Integration of Construction Information." Department of Surveying and Information Technology Institute, University of Salford, Salford.

Aouad, G., Child, T., Marir, F., and Brandon, P. (1997). "Open Systems for Construction (OSCON), draft industry report." Department of Surveying, University of Salford, Salford.

Cunningham, J.J. and Dixon, J.R. (1988). "Designing with Features: The Origin of Features," ASME Computers in Engineering Conference, San Francisco, CA, Aug. 1988, 237-243.

Darwiche, A., Levitt, R., and Hayes-Roth, B. (1989). "OARPLAN: Generating Project Plans by Reasoning about Objects, Actions and Resources." AI EDAM, 2(3), 169-181.

Fischer, M. (1991). "Constructibility Input to Preliminary Design of Reinforced Concrete Structures." Technical Report 64, Center for Integrated Facility Engineering, Stanford.

Froese, T. M. (1992). "Integrated Computer-Aided Project Management Through Standard Object-Oriented Models." PhD Thesis, Stanford University, Stanford.

Hanna, A. S., Willenbrock, J.H., and Sanvido, V. E. (1992). "Knowledge Acquisition and Development for Formwork Selection System." Journal of Construction Engineering and Management, ASCE, 118(1), 179-198.

International Alliance of Interoperability (IAI) (2001). "IFC 2x Extension Modeling Guide", Available from http://www.iai.org.uk.

Laitinen, J. (1998). "Model Based Construction Process Management." PhD Thesis, Royal Institute of Technology. Stockholm, Sweden.

Staub-French, S., and Fischer, M. (2001). "Industrial Case Study of Electronic Design, Cost, and Schedule Integration." Technical Report 122, Center for Integrated Facility Engineering, Stanford.

Staub-French, Sheryl (2002) "Feature-Driven Activity-based Cost Estimating." PhD Thesis, Stanford University, Stanford.

Thomas, H.R. and Zavrski, I. (1999). "Construction Baseline Productivity: Theory and Practice." Journal of Construction

Engineering and Management, American Society of Civil Engineers, 125(5), 295-303.

Timberline Software Company (2001). Precision Estimating Extended and CAD Integrator, Users Documentation, Beaverton, Oregon.

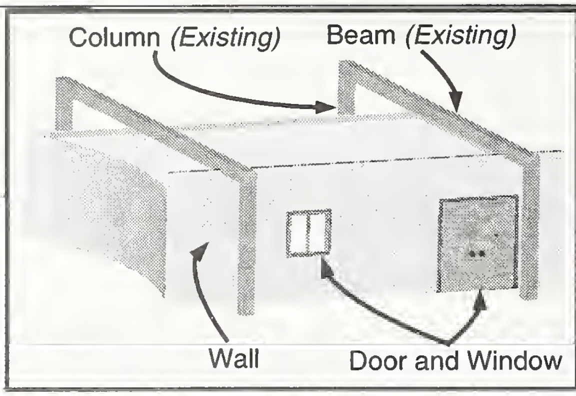


Figure 1. Building components in the office project case study.

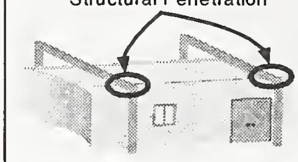
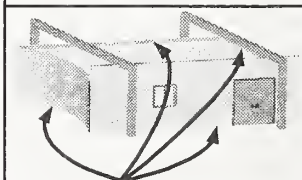
Relevant Design Conditions	*Estimator's Rationale*
 Curved Wall	Reduces the crew productivity rate for all the "Install Metal Studs" activities.
 Wall Height	Uses Rolling Scaffolding and reduces crew productiivty in the "Install Metal Studs" activity if the wall height is between 9' - 13'.
Structural Penetration 	Adds activity "Frame Penetration" to account for the additional labor costs for the unusual framing condition. Adds activity "Apply Caulk" if the intersected wall is fire-rated.
 Component Similarity	If most of the walls have the same height, increases the crew productivity for the "Install Metal Studs" activity by 20%.

Figure 2. Estimators' rationale for adjusting the activities, resources, and resources' productivity rates to reflect the cost impact of specific design conditions in the cost estimate.

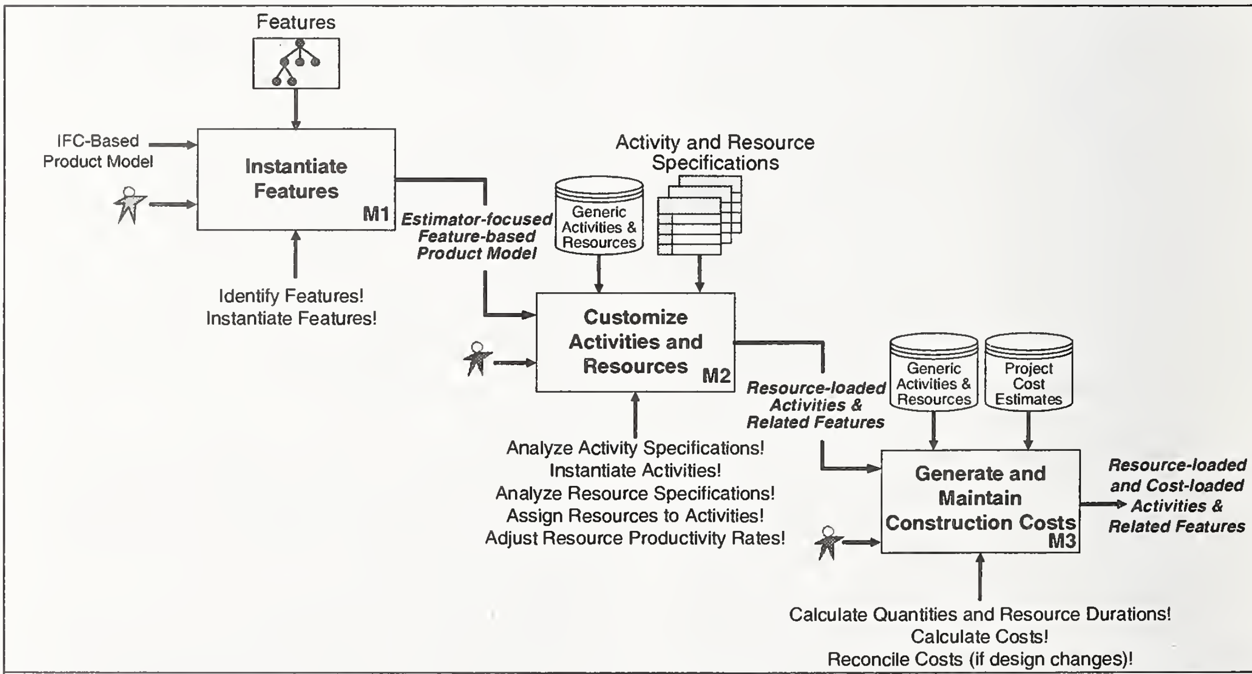


Figure 3. The three steps of the activity-based cost estimating process to (1) create an estimator-focused feature-based product model, (2) customize the activities and resources based on the estimator's rationale in Activity and Resource Specifications, and (3) create resource-loaded and cost-loaded activities that are related to the features in the estimator-focused feature-based product model.

Construction Zone Generation Mechanisms and Applications

Ragip Akbas

Stanford University
Center for Integrated Facility Engineering
M:4020, Stanford, CA, USA 94305-4020

rakbas@stanford.edu

Martin Fischer

Stanford University
Construction Engineering and Management Program,
Terman Engineering Center, Room 292
Stanford, CA, USA 94305-4020
fischer@stanford.edu

Abstract. Construction zones are geometric units of work in the construction planning process. Using construction parameters and geometric properties, construction zone generation mechanisms change the level of detail of the 3D geometry and corresponding activities of the planned project. We have defined and implemented a set of mechanisms for construction zone generation. *Decomposition* provides various ways to generate detailed zones, while *aggregation* generates combined zones. To support these mechanisms, we have formalized a geometric representation based on triangular meshes that can represent the activity workflow, project spatial hierarchy, local variations in productivity, and activity state information on a component at any given time. This paper summarizes the mechanisms and describes various planning/scheduling applications of these mechanisms, such as rapid plan evaluation, resource and workflow balancing in 3D, and association of schedules at multiple levels of detail. Our contributions are the formalization of these zoning mechanisms, representation and algorithms to support these mechanisms, and planning/balancing methods utilizing the mechanisms.

Keywords: construction planning; geometric algorithms; construction zones; line-of-balance techniques

1. INTRODUCTION

Today, it is difficult for construction superintendents and schedulers to consider and communicate alternative plans even with the existence of 3D models. There is also a gap between the techniques for high-level planning/scheduling and detailed operations level planning from research and practical perspectives. We designed construction zone generation mechanisms to address these problems.

Where and when an activity is ongoing is crucial knowledge that relates different sources of information. Scheduling at a more detailed level can decrease activity duration and increase the reliability in the construction schedules. Relating high-level and detailed schedules can prevent common inconsistencies between schedules at different granularity.

One of our goals is to make effective use of geometric models, extending existing techniques in computer graphics to build a general computational structure for construction planning. We want to make the geometry details transparent to the planner, capturing the reasons for planning as parameters in zoning mechanisms. Then, changing these parameters, planners can easily try different scenarios with different crews and methods and different workflow parameters.

1.1 Previous Work

We define construction zones as the geometric units of work in the construction planning process. A construction zone can be a part of a component, a component, a group of components or a volume.

We refer to the process of changing the level of detail of the building geometry and activities as *zone generation*. Previous zone generation approaches are mostly for automated planning and use user-defined volumetric elements. Their purpose is to simplify the automated planning process by creating sub-networks of activities and constraints within zones and combining them to generate the final plan.

Thabet and Beliveau [5] define zones for aggregating components using a hierarchy of boxes. The ZonePlanner research [6] uses volumetric grids to search for an optimal zoning plan. These systems do not address crucial knowledge of planners related to workflow, arrangement of work areas, and changes in production rates due to various factors. None of these systems analyze and make use of complex geometry. They suffer from the necessity to define building blocks for each case a priori and from the combinatorial complexity to try all possible merging alternatives.

Riley and Sanvido [3], in their research for space planning in multistory buildings, define work-area patterns to describe the directions for and locations of units of work completed for different activities and materials. However, they don't represent the patterns computationally. We provide a computational method for supporting similar patterns of construction on 3D models.

4D CAD, in its simplest form, is the ability to visualize the state of construction at any given project time. It uses the association of the building elements with the activities required to construct them. Therefore, every component has time attributes. We extend that to every polygon on the geometry. Therefore, every location on the geometry can support different time parameters.

1.2 Terminology

We represent a building element by association of semantic product information with geometry, similar to many product modeling approaches [1]. We use the term building element synonymously with an assembly, e.g., a stud wall, concrete slab, roof. Building elements or assemblies are organized in a spatial or organizational hierarchy, a tree, where only the leaf nodes can contain the actual components (Figure 1). An assembly is composed of one or more of its components or constituents. For example, components for a stud wall assembly are studs, dry wall, insulation, etc. Each component can have different partitions representing zones, the smallest possible partition being the unit size of that component. Each assembly has a type, e.g., floor slab. By the same token, each activity has an *activity type*, e.g., pour concrete.

There is no single correct spatial level of detail for building elements. For example, either the whole floor slab or part of this slab enclosed by a room can be a building element with different granularity. Similarly, the activity required to construct it can be for the whole slab or the slab bounded by the room.

We distinguish between three main component categories for zone generation purposes: *prefabricated*, *fixed size site-assembly* and *continuous site-assembly*. Prefabricated components are installed as a single step. Therefore, the component and the assembly are the same, and they cannot be partitioned. Fixed-size site-assembly components require on-site installation, are formed of well-defined units of work, and one or more different activities are necessary for their construction, e.g. stud walls or roofing. Continuous site-assembly components are either not formed of unit size elements or the elements are negligible in size, such as reinforced-concrete components.

In the rest of the paper we will first explain the zoning mechanisms and the required geometric representation for them. Then, we will focus on different applications using these mechanisms in areas where traditional planning and scheduling techniques are limited.

2. ZONE GENERATION APPROACH

2.1 Mechanisms

We have abstracted the common level of detail changes during the construction planning/scheduling process using a well-defined set of zoning mechanisms, which act on components and activities. Zoning mechanisms essentially describe planning and geometric parameters for how activities are installed. They are powerful in replicating the construction planner/scheduler's intent and support what-if analyses and visualization. They are also designed to be general and extensible. We have two main categories of mechanisms for zone generation: *decomposition* and *aggregation*. An intuitive way to think about these mechanisms is: *decomposition* creates and maintains zones formed by a part of a component, *aggregation* creates and maintains zones consisting of a group of components.

2.1.1 Decomposition

This mechanism generates detailed construction zones on components and detailed activities for the zones, thereby increasing the level of detail of the existing models. The operation for decomposition is Decompose(*c, a, t, D, F*), where
c: component to be decomposed,
a: activity to be decomposed,
t: type of decomposition, one of (grid, production),
D: direction set, which contains, in the case of grid decomposition, a direction vector and amount of decomposition in each direction. For production-based decomposition, the direction set is the starting location and a set of unit direction vectors at specified locations,
F: factor set, which contains the shape factor functions and parameters and time-space functions as explained in section 2.3.

After the decomposition, the application generates a set of *subcomponents* and *subactivities* and links them to each other. The geometric representation for each subcomponent is a partition of the geometry for the original component. Each subactivity is a result of splitting the original activity.

2.1.2 Aggregation

This mechanism groups components and sets the start and finish dates by associating detailed activities with the new group of components. Its types range from organizing the components within a region to ordering the activities necessary for their construction. It supports, for example, ordering of exterior walls given the starting location and direction.

The operation is *Aggregate*(*node, a, t, D, F*), where *node* is a node in the product hierarchy, *t* is the type of aggregation, *D* is the direction set, and *F* is the factor set. Note that *Aggregate* is different from the *Union* operator from solid modeling, in that its purpose is not to combine the geometry of individual components. Furthermore, *Aggregate* is not the inverse operator of *Decompose*.

2.2 Geometric representation for the mechanisms

Since the zoning mechanisms act on the geometry, geometric representation and algorithms are an important part of this research. We use triangle meshes for the geometric representation. Geometrically, a triangle mesh is a surface consisting of triangular faces pasted together along their edges. The mesh geometry can be denoted by a tuple $M = (K, V)$, where K specifies the connectivity of the mesh (the adjacency of the vertices, edges, and faces) and V is the set of vertex locations defining the shape of the mesh in $\mathbf{R}^3$. The vertex locations and the connectivity for the mesh can be acquired from any 3D CAD or modeling application. We can also build the connectivity for the faces using common computer graphics techniques.

Triangle mesh is a common and well-researched representation. However, as it is, it lacks the construction domain information needed to support the zoning mechanisms. We have extended the mesh representation to a tuple $M = (K, V, D, S)$, where D is local direction of workflow and production modifier at any location on a component, and S is the state of the component location as a function of time. D is associated with a (vertex, activity) pair, while the state S is associated with the faces (Figure 2c). These main extensions are as follows:

- Every vertex on the mesh contains a unit direction vector for each related activity to describe the local direction of workflow at that location.
- Every vertex on the mesh contains a productivity modifier for all related activities, $Pk \in [0,1]$. Originally, modifiers on all vertices are 1, meaning the activity production rate at the given location is equal to the production capacity.
- Every triangle contains a *local activity state* to represent the installation status at any given time. We created an *activity state machine* to unambiguously calculate the local activity states as a result of the zone generation mechanisms.

2.3 Production Rate Modifiers

The production rate is not always constant for activities anywhere on a component. We support variable local production rates using the productivity modifiers using the parameters stored with the geometry. The modifier can be a combination of the effects of the component's shape on the production rate (*shape factors*), effects of other ongoing activities (*time-space factors*), or other effects that we did not consider so far in our research (learning curve, availability of resources, etc.).

The application calculates the production rate value using the user-supplied functions and the shape representation. Shape functions are dependent on geometric parameters including height, distance to edge, and slope. Schedulers declare how production rates are affected by these parameters so that the software can calculate the local production rates. Time-space factors can decrease the production rate if two close activities are installed at the same time in the schedule. They can also stop an activity if a prerequisite component is not ready.

2.4 Spatial Structures

We use various spatial structures to aid the mechanisms. A *region* is a bounded volume, which encloses all the components that fall into it. Examples of regions for construction projects are a floor, a room, or a quadrant. We support a hierarchical region structure that is user defined and unique for a project. Other spatial structures we use are *octrees* for spatial search and *grids* for arrangements of some mechanisms.

3. APPLICATIONS

We now will use the defined mechanisms together with the geometric representation to show approaches to common construction planning-scheduling problems. Starting with plan evaluation on the simplest, single building element case, we will extend the application to multiple building elements at variable levels of detail. Building on the common line-of-balance convention of continuous utilization of resources, we will describe a hierarchical 3D scheduling approach using the zoning mechanisms. Then, we will consider extensions to the architecture, such as keeping an operational schedule synchronized with the master schedule, and defining construction method-specific zones using the mechanisms.

3.1 Rapid plan evaluation

This application is a direct consequence of the described mechanisms and the representation. Users can consider planning alternatives quickly by changing the direction of workflow, starting location, production rate and the sequence for the activities. The application uses the parameters in the mechanisms to traverse the geometry and generate the states for any location at any time.
We have a prototype application that allows the user to enter the parameters for the mechanisms, which become input for the actual mechanisms. Let's say the user wants to plan the pour for the slab in Figure 1. The user interactively defines how it will be constructed in the application, in this case specifying the direction of flow vector and amount (Figure 2a), which in turn generates the following mechanism:

Decompose(Floor2slab,Pour_Floor2slab, *grid*, *D*, *F*)

Note that if the user changes the direction of the activity or the amount of decomposition, only the directional information, *D*, changes from the initial parameters. The user can immediately visualize the effect of this change. If the user wanted to perform this plan evaluation manually, he would need to (1) modify the 3D model to split the geometry, (2) split the construction activities into subactivities, and (3) re-link them. Any other change requires the repetition of this tedious process.

For the columns over the slabs, the zones are groups of columns (Figure 2b). The mechanism is:

Aggregate(F2Columns, Rebar_F2Columns, *production*, (A-1, (-1,0,0)), *F*)

Obviously, construction projects are composed of many activities and components, and considering activities in isolation is not very valuable. The user can also decompose the same slab for the other associated activities, such as *rebar* and *form* by applying *decompose* mechanisms for each. In that case, different activities on the same component should follow each other without interference. The *activity state machine* guarantees that the activities do not violate their predefined ordering, i.e., the prerequisite activity is always installed before, and there are no two activities installed at the same location at the same time. The mechanisms are also applicable to more complex structures. Figure 3 shows production-based decomposition applied to an exterior enclosure element. The user specifies the starting location, direction of workflow, and effect of shape on production rate for different activity types required to construct the enclosure.

The following application focuses on the interactions of multiple activities on multiple building elements considering resources, providing an alternative scheduling approach when 3D models are available.

3.2 3D schedule balancing

Traditionally, line-of-balance (LOB) or linear scheduling techniques are used for activities following each other for linear or repetitive work, allowing planners to balance workflow by adjusting the production rates and start dates for activities [4]. They help plan for continuous utilization of resources. However, these methods are limited to fixed single level work areas corresponding to the flow. Activities in linear scheduling techniques should either be repetitive units or proceed linearly in a single axis-aligned direction - horizontal or vertical. To plan using LOB, the project is divided into planning units usually as large work areas. Although multiple activities can concurrently act on the same work area, LOB conservatively limits a single activity in the same work area at a time.

Our application extends these valuable techniques to many types of construction activities using the spatial and planning information and benefiting from the same zoning mechanisms explained in the previous application. It adds the ability to schedule at multiple levels and supports multiple activities on a component and complex shapes. The balancing process can range from single element/single region to hierarchical/multiple regions.

3.3 Maintaining master and operational schedules synchronously

The previous application examples ignored maintaining the level-of-detail relationships with the activities in the existing master schedule. Master schedules are the means for tracking resources and activities for construction projects. Detailed planning with zoning mechanisms and actual detailed site data adds valuable input to the schedule. Project scheduling software [2] provide hammock or summary activity types to summarize a group of activities. However, this relationship is error-prone and limited.

We associate the original activities with the activities resulting from the zoning mechanisms through the associated project locations. Whenever the activity date and duration are modified at a more detailed level, the activities in the master schedule can be updated. Similarly, users can propagate the effect of changes in the master schedule to the detailed schedules. This also allows the construction process visualization at two levels of detail.

3.4 Method changes on component

In all of the previous examples, the generated zones correspond to some planning units of work. Zones can also support changes in construction methods on a single component. Construction methods can vary because of geometric factors such as height and slope or existence of nearby components, such as scaffolding. We support ways to generate separate zones on building elements in such cases.

4. CONCLUSIONS

In this paper, we have briefly described construction zoning mechanisms and explained interactive planning/scheduling applications. We are validating the research by replicating the manual operation plans on several projects via the zoning mechanisms. The tested activity types include concrete pour schedules, exterior enclosure and roofing.

Zoning mechanisms assume an activity proceeds in a connected way and make a surface approximation for component geometry. Additionally, we use linear shape and time-space functions.

These mechanisms provide the possibility of many other types of visualization, analysis and optimization. Another important use is data collection by reverse-engineering the actual construction dates to infer the planning parameters. A space planning methodology can be adopted for the 3D balancing application to check for other types of spatial conflicts. Furthermore, functions to generate the temporary structures such as laydown, scaffolding and formwork on the fly can increase the value of this research.

5. REFERENCES:

1. IAI. 2001. *Industry Foundation Classes - Release 2.x Specifications.*

2. Primavera. 2000. *Primavera Project Planner*. Primavera Systems, Inc., Bala Cynwyd, PA.

3. Riley, D. R. and Sanvido, V. E. 1995. 'Patterns of Construction-Space in Multistory Buildings', *Journal of Construction Engineering and Management,* Vol. 121, no. 4, pp. 464-473.

4. Stradal, O. and Cacha, J. 1982. 'Time Space Scheduling Method.' *Journal of Construction Division*, Vol. 108, no. 3, pp. 445-457.

5. Thabet, W. Y. and Beliveau, Y. J. 1997. 'SCaRC: Space-Constrained Resource-Constrained scheduling system', *Journal of Computing in Civil Engineering*, Vol. 2, no. 1, pp. 48-59.

6. Winstanley, G. and Hoshi, K. 1993. 'Activity Aggregation in Model-Based AI Planning Systems', *I. Journal of Artificial Intelligence in Engineering Design, Analysis and Manufacture (AI EDAM)*, Vol. 7, No.3, pp 209-228.

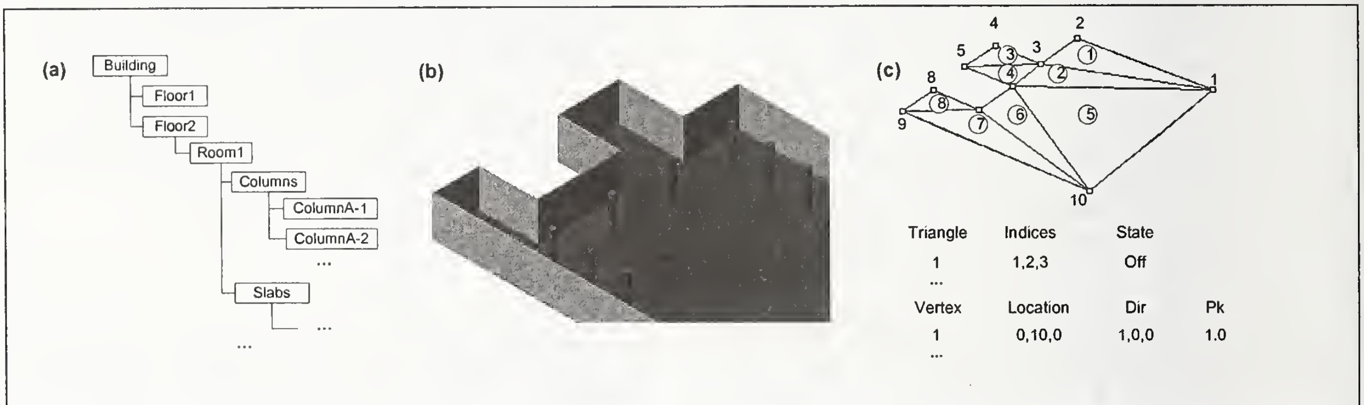


Figure 1. Representation basics for the sample model. (a) Partial product composition hierarchy. (b) Partial 3D isometric view. (c) Triangle mesh representation for the slab including the extensions.

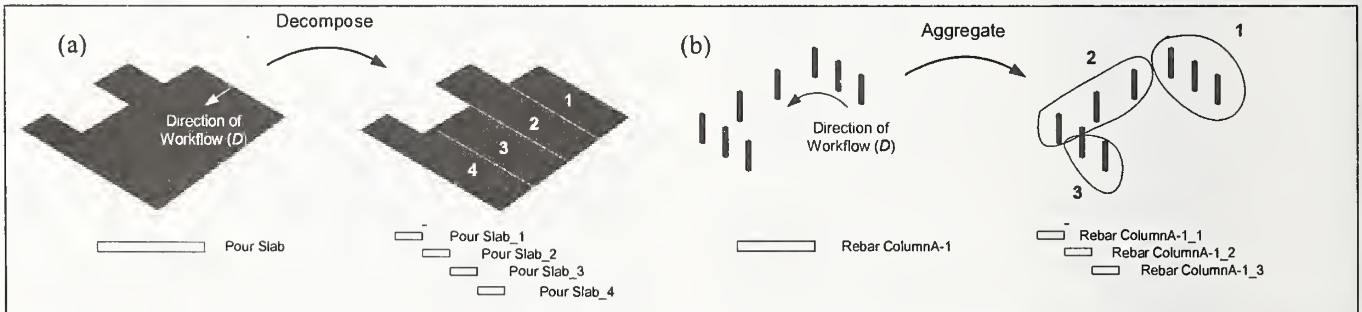


Figure 2. Examples of mechanisms. Decomposition splits the geometric representation to correspond to the units of work for *pour_slab*. Aggregation groups the concrete columns for *install_rebar* in this example.

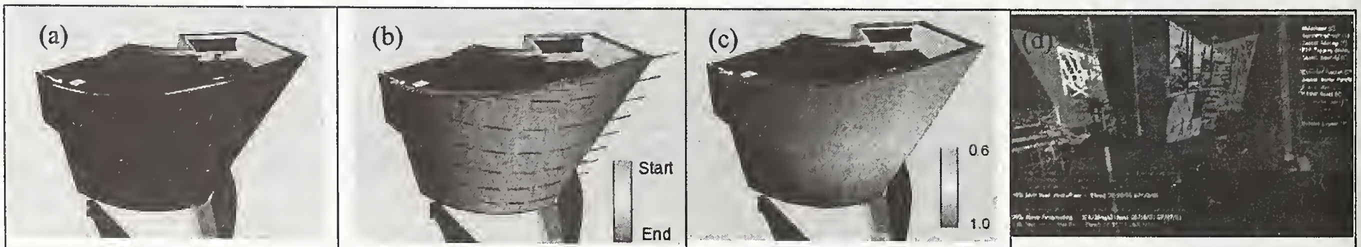


Figure 3. Zoning mechanisms applied to a complex exterior enclosure surface. (a) Original model. (b) Direction of workflow interpolated over the surface visualized by arrows and colors. (c) Height factor calculated on the surface. (d) Snapshot of 4D model after the mechanism is applied.

ASSEMBLY + DISASSEMBLY OF INTERIOR WALL

by

Dipl-Eng. Albrecht Hanser

Technical University Munich, chair of building realization + -informatics, Prof. T. Bock Arcisstrasse 21, D 80333 München,
albrecht.hanser@bri.ar.tum.de

Abstract: Flexible extension walls - watched on a long-term-use - have lower costs and are more economical and especially more variable compared to convenient post-and- beam-structures. Because room requirements – mainly in office and administration building – change faster and faster, an extensive potential on future markets for these technologies exists

Keywords: Flexibility, Economic Efficiency, Non-supporting removable wall-unity, Media- and electric wiring, Put-and-slide-technology, Relocate systems, Cost, Different Elements,

1 INTRODUCTION

Flexibility

A higher and higher flexibility of the floor plan is expected in office- and commercial building as in house building. The change of users or function often results in the wish redesigning the existing spatial pattern. A basic floor plan flexibility often is possible and practiced by dividing the load-bearing building surface from non-structural separation walls and ceilings. A well known system of this type of walls is a wooden or metal post-and-beam-structure with mineral fibre filling and plasterboard panelling.

By using this building system it is difficult to realize variations of different room structures and subdivisions by self building. In fact they are cheap to produce and to install but in the end they are much fuss in redesigning, few flexible and actually not reusable. In addition there are already reusable, demountable partition systems available. However they usually are to expensive and to much fuss to construct.

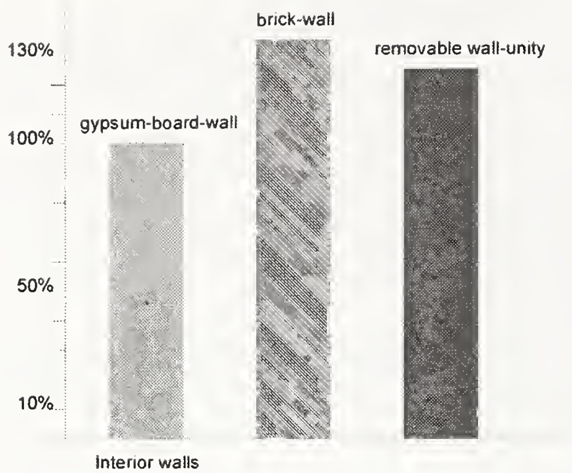
Economic Efficiency

All partners in construction business know exactly what construction cost mean. Very often we think and act only in this kind of category.

The reason for this understanding is the existence of two different budgets; one for construction and a second for consequential costs during life cycle, which are even managed by two teams. Feedback between the two will almost not occur. So at the end in a life cycle of a building, nobody is able to find out the real cost situation which has been followed after construction time, of a product incorporated in the building. Even what the savings could have been, we cannot find out, in case we would have constructed in prefabricated and modular systems.

Indeed thinking in a complex overall view, we are realizing where the advantages are for prefabricated and relocate-partitions:

- Connections with its components, seperations in floor- and ceiling-areas are close in cost of these construction types.
- The finishes of partition-walls, out of gypsum board, assembled on site, have not this standard of quality as industrial prefabricated partitions. Prefabricated partitions need no new coatings for 30 years, as conventional gypsum boards need it all 5 years.



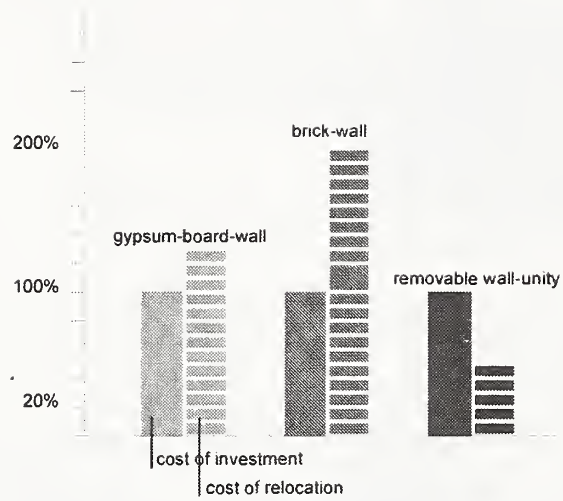


Figure 1: interior walls: cost of investment and relocation (1)

- In case prefabricated partitions get relocated, the difference in investment cost becomes obvious. A prefabricated partition occurs 40% of its former construction cost, while a conventional partition, which has to be demolished and reconstructed almost completely new, occurs 220% fit. It is beyond any doubt, the payout-return depends on construction type of prefabrication and the amount of relocations.
- Recovery and cleaning up belong to an ecological and oeconomical overview; in this context relocate systems offer advantages. Savings of material and avoidance of waste have no exceeding fees for supervised dumps. These cost have been augmented constantly the last years.

2 AIM OF THE PROJECT

It was valid to develop a non-supporting removable wall-unity, which allows adaptable, fast and disassembling room-splitting.
Next to the standard components we have to develop elements for the connection at the existing support structure, elements for passages and doors, for windows on the inside and other glazing structures.

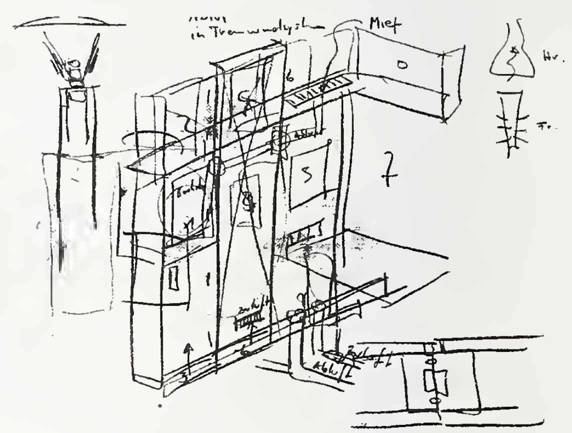


Figure 2: First sketch (4)

Media- and electric wiring are components of the removable wall-unity. The elements are delivered prefabricated and manageable on the construction site.
The connection to the available supportstructure is created with a streching mechanism (Justix); the coupling of the elements and also the management of media- and electric among each other works with a put-and-slide-technology.

3 DEVELOPMENT

Demands to the system
- "More than a wall", which means integrated media, such as TV, phone, flat screen, switches, bus, safe, mini-bar, shelf-systems, cupboards, etc completely connected up;
- Demountable and reusable and without damaging the existing solid construction;

- With finish-surface on both sides, in complete lightweight construction;
- For two persons to handle and to install
- Simple construction and long-term use

- Located in price between the plaster board/ post-and-beam-structures and the conventional wall systems.

"More than a wall"; beside separating for offices, administration or housing the wall unit is supposed to serve as room divider semi high screen wall or fair stand.
Put on wheels it can even become a mobile office wall.

One of our aims is to see the wall not only as a separating element between two areas, but to include additional functions such as screens lights, mini-bars, shelf-boards, integrated air-condition, etc; even PC or printer can be hidden in the wall;

4 ELEMENTS

Surface finished, easy-to-handle modules are basic for interior wall surfaces.
Elements are:

- Room-high elements (size depending on height of the ceiling, standard sizes)
- Door high elements (height 210 cm)
- End elements above door high elements (size depending on height of the ceiling)
- Elements for easy zoning, separating areas (height 105 cm)
- Special elements

5 CONSTRUCTION

The substructure for the single module is supposed to be a circulating frame. Panel materials for wall surfaces like wooden panels, plasterboards, glass etc. serve as frame bracing.
Wooden profiles but also special formed plates can be the frame material.
The connection to the load bearing wall construction and to the ceiling is designed as edging board.

The gap between the bottom of the wall an the floor is covered by a clipped baseboard. A dovetail-similar joint with integrated plugs for power and data supply connects the single elements among one another and to the load bearing walls.
The single modules therefore are coupled by

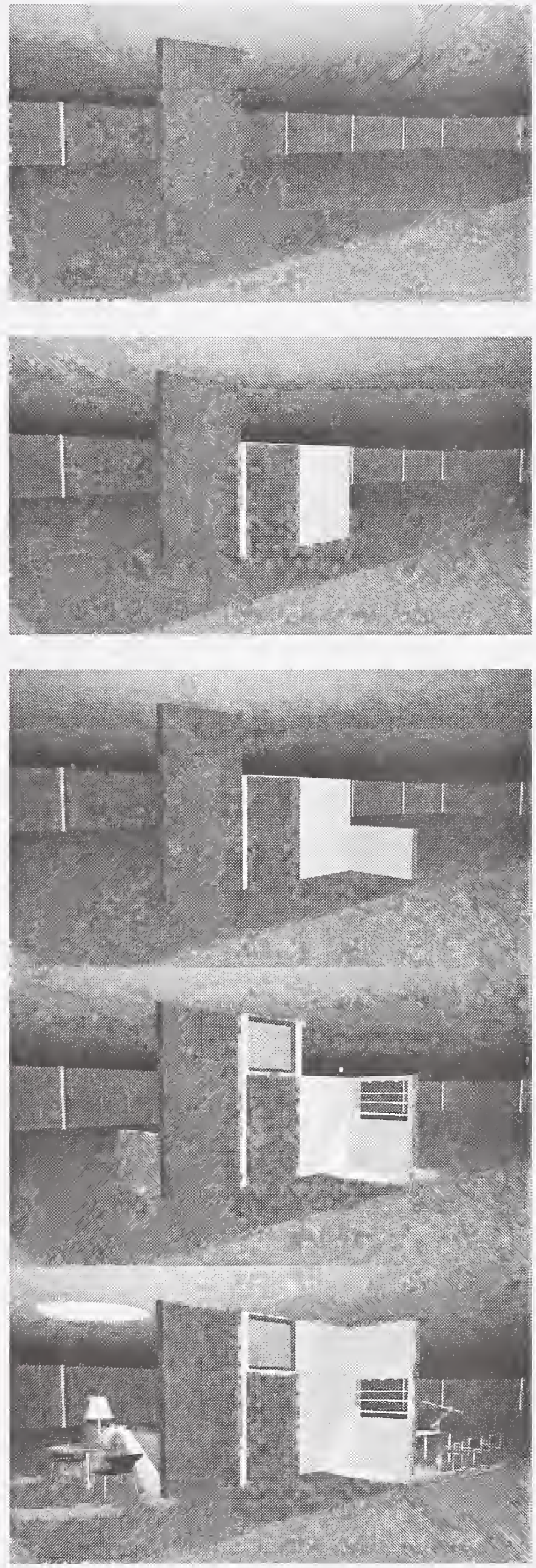


Figure 3: Simulation of the elements (4) fitting an pushing the connecting elements

The single modules therefore are coupledby fitting an pushing the connecting elements together. The stack design is variable.

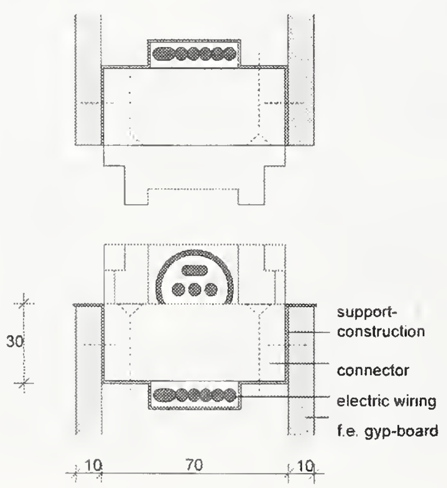
Door high elements and end elements are connected in the same way. A special processing of the frame profiles makes it possible to carry the prefabricated electronic wires through the element.

The contact from element to element gets possible through the just mentioned connecting modules. The necessary fitting in to surround construction could be succeeded by vertical adjustable foot construction. There are two for each wall panel. After connecting the elements the foot construction is screwed up and pressed against a connecting profile.

The rising gap at the bottom of the wall can be used for supplementary wires – especially for wire systems, that can only be added restrictively (phone, networks). A clipping profile as baseboard covers the open bottom area. There must be sound and fire technical measures on the surface, in the stack area as in the base and ceiling area.

6 FABRICATING PROTOTYPES

There are running parallel extensive tests with prototype wall modules to transfer the gained knowledge practical.



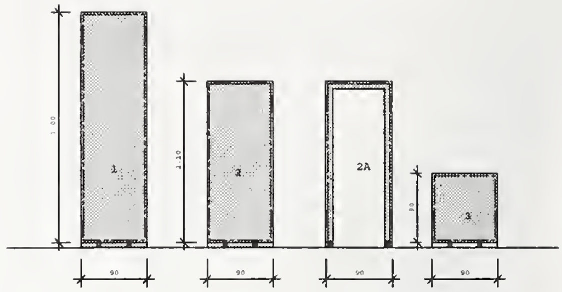


Figure 4: The elements; dimension: 90cm wide, 300cm, 210cm, 90cm high; the connecting system (4)

7 CONCLUSION

The system shown in this paper will find place between the low cost but not reusable gypsum board partition-walls and the industrial prefabricated partitions, which are very expensive. The combination between a prefabricated high tech system, affordable and economical, usable as flexible elements for different tasks are the main advantages of the new system.

REFERENCES

(1) Studiengemeinschaft Fertigbau (1987). Trennwandsysteme

(2) Tichelmann / Pfau (2000). Entwicklungswandel Wohnungsbau, Vieweg, Wiesbaden

(3) Diverse Herstellerunterlagen: Lindner, Knauf, Item, Dorma, Häfele, Strähle, Hettich, u.a.

(4) all figures: Technical University Munich, chair of building realization; exept figure 1 which belongs to reference (1)

SESSION 6

FIELD SENSOR DATA AND CONSTRUCTION PROCESS INTEGRATION PROTOCOLS

STANDARDIZATION OF DATA FLOWS ON EARTHWORKS AND ROAD PAVEMENT SITES USING INFORMATION SYSTEMS

by

François Peyret

ABSTRACT: The work described in this paper is the follow-up of the first analysis presented in the previous ISARC on the same theme. The paper start by recalling the background with is the ISO TC-127 WG2 standardization initiative. Then he applies a basic logical model of site information system for civil-engineering sites to various existing or under development such European systems (CIRC and OSYRIS) to identify and specify the main categories of information flows that are worth to be standardized. The paper ends with some considerations about these categories and the way to continue the work.

KEYWORDS: standardization, site information systems, civil-engineering, computer integrated construction, data flows.

1 INTRODUCTION

Still much time and money are lost all along and between the various phases of construction, due to the lack of information and the very poor quality of the management of this information, making very small use of the new available Information Technology. Clearly, the new challenge in the construction world is now the information management.

Due to the presence, on more and more work sites, of novel operator-aiding systems using smart positioning and processing systems, people have become aware of the value of all the relevant data that can be collected from such systems, for instance for quality assessment and relationships with the road owner. These data, that need to be exchanged between different information and processing systems are really valuable only if they are standardized

2 BACKGROUND

To address this issue, the ISO TC-127 standardization committee launched a new working group, WG 2, focused on the scope of *Work site data controlled earth-moving operation* in October 2000. This group, gathering experts from Japan, USA, France, Germany, Sweden and Italy, already met 3 times, in Tokyo, Bologna and Denver and the main decisions which came out from the discussions are the followings:

- to focus upon road construction works, including earth moving operations and pavement construction,
- to start with the standardization of the definition and content of the information exchanged, then to study the format, then the protocols,
- to stay close to the site, and more precisely to the machines itself, that is to say not to study, in a first step, the higher level exchange of information between road authorities and the contractors.

To structure its work, the WG decided to subdivide it into 3 projects: *Terminology*, *System Architecture* and *Data Dictionary*, the first two ones being under the responsibility of Japanese experts, the third one of the French experts, co-operating with the American ones.

Terminology should set up the basis of common understanding for all the work, *System Architecture* should define how the data will be handled in the various exchanges and *Data Dictionary* should define what data should be standardize and propose a common definition and representation of them.

Site Robotics and Positioning Team, LCPC (Laboratoire Central des Ponts et Chaussées), Bouguenais, France, francois.peyret@lcpc.fr

This paper presents mainly the work achieved by LCPC, composing the French delegation, about data identification and clustering for preparing the work of establishing the relevant data dictionary to be included in the tentative standard.

3 THE LOGICAL MODEL

The paper presented at last ISARC in Warsaw [1] already presented differences between logical and physical modeling and explained why we proposed to start the standardization discussions at the logical level. This paper also introduced the important concept of 'road product model'. The core part of the paper was the proposal of a logical model of a *Site Information System* which was based upon the analysis of the functional design made for the OSYRIS project [2].

A 'system' is a group of functions providing a 'service' (can be an action but generally a piece of information) as answer to a 'demand' (always a piece of information). Each system can be decomposed into several sub-systems, thus giving a fractal structure to the representation.

The logical model of our system is presented in Figure 1, is composed of four main sub-systems which play well-separated roles in the global process: *Temporary Data Storage*, *Data Collection and Processing*, *Assist For Operation* and *Work Execute* systems.

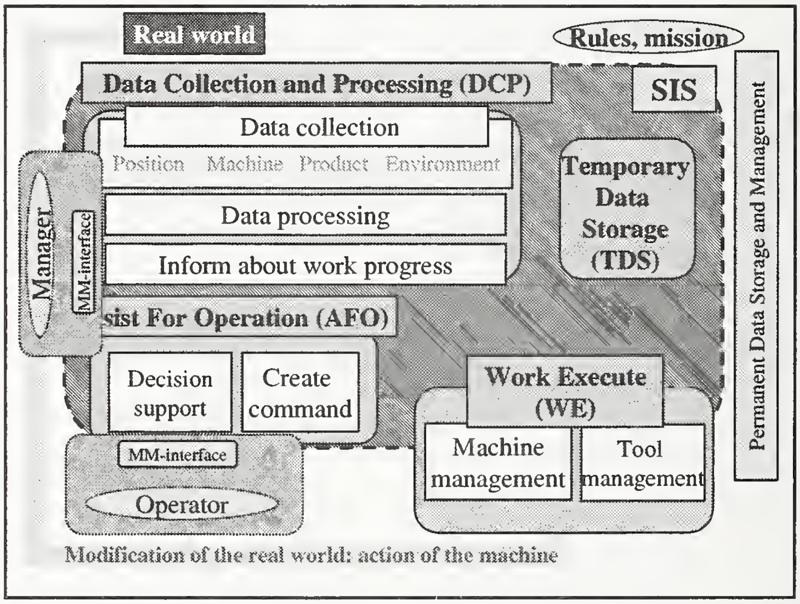


Figure 1: The basic logical model

- *The Site Information System (SIS)* (formerly called *Data Operation System*) means the whole system itself which is managing information to improve the quality of the work. The SIS can be physically distributed at various places, on the machines. It has generally one 'office' segment and one 'site' segment.
- The *Work Execute (WE)* sub-system is the part of the SIS who is managing the machine and/or the tool. It might not exist when the operator is steering manually.
- The *Assist For Operation (AFO)* sub-system, which can also be called *Execution Management* sub-system, is the sub-system which is making the decisions for the execution of the work, with or without the help of the operator.
- The *Temporary Data Storage (TDS)* sub-system is in charge of storing all the temporary data that important for the quality assessment, before transmitting it to the external 'Permanent Data Storage and Management System'.
- The *Data Collection and Processing (DCP)* sub-system, is a composite and intermediate system which is in charge of acquiring the necessary data from the Real World or from the TDS and in charge of processing them for the other sub-systems. It also prepares the data for visualization if necessary.

The human beings participating to the execution, grouped into two classes: *Manager* and *Operator*, are considered external to the SIS, thus contributing essentially to the process. The machine itself, that is to say the piece of equipment composed mainly of hardware but also of software, under the form it is provided by the manufacturer, is considered also external to the SIS.

The data flows that are in the scope of the standardization work are those exchanged between: SIS and machinery, SIS and measurement instruments, SIS and contractor or consultant office.

4 DATA EXCHANGED BETWEEN THE VARIOUS SUB-SYSTEMS.

4.1 Analysis Of Existing Systems

To support our analysis, we studied carefully the data flows existing in several SIS already developed or under development. The SIS we analyzed are dedicated to road pavement construction, but we make the assumption that the results would be valid for earth moving sites too, with maybe some minor differences.

First, we considered the *Computer Integrated Road Construction* [3] products, that is to say the CIRCOM system, dedicated to the rollers, and the CIRPAV system, dedicated to the pavers.

CIRCOM

CIRCOM [4] is an operator-aiding system which provides, through an accurate and ergonomic display (colored map), to the operator the position of its machines, its speed and the already achieved number of passes. The positioning component, called CIRCOM POS, is a quite sophisticated one, using various sensors, in particular a high-precision GPS. The preparation of the mission and the transmission of the design data are made on a computer called CIRC GS (for Ground Station) and the on-board processing is done by the CIRCOM OB (for On-Board) computer.

The CIRCOM data flow analysis is presented on Figure 2.

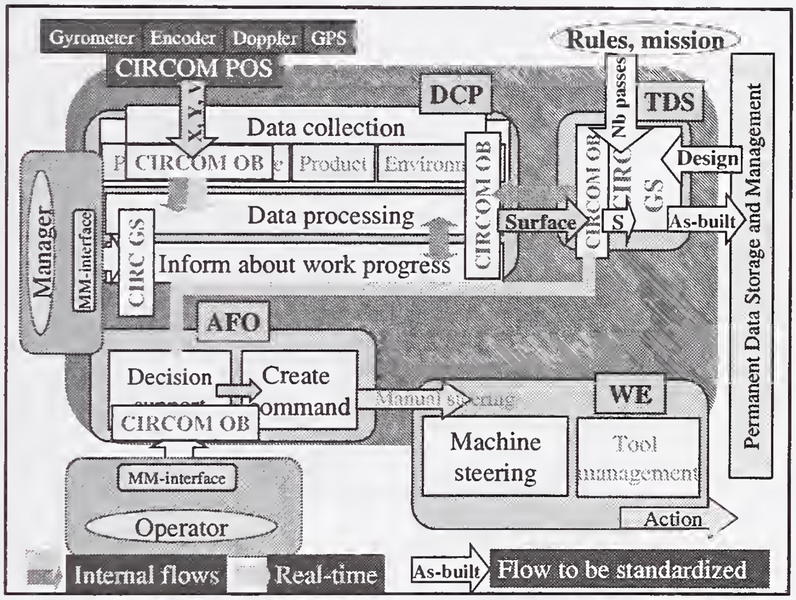


Figure 2: The CIRCOM data flows

The first thing to remark about the CIRCOM SIS is that the Work Execute sub-system (WE) doesn't exist as the steering of the machine remains manual, or can be considered as external.

The second point to notice is that the physical 'components' of the system are distributed among the various logical sub-systems, meaning that a physical component provides generally different kind of services, from a logical point of view (e.g. temporary storage, intermediate data processing and display to the operator for the on-board computer).

The data flows that are activated by the operation of the system are of different kinds. Some are 'internal', meaning that they should remain proprietary and thus, are not to be considered for standardization. Some are real-time and some are not. The important data flows for standardization are those which are exchanged either between the SIS and external systems or between sub-systems, when it is necessary. This has to be discussed case by case.

1. Data exchanged between the SIS and external systems:

- *design* data (description of the road geometry);
- *mission* data (target number of passes and target speed)
- *as-built* data (compaction map showing the actual achieve number of passes on the road);
- *positioning* data (plane position and speed), from the positioning component, considered external.

2. Data exchanged between sub-systems:
we took into consideration what we called the *achieved surface* data, that is to say the 'just achieved' work of the roller that might be interesting to transmit and to use in real-time. This is justified for this kind of system because these data have to be exchanged between the different compactors and they might be equipped with similar systems from different manufacturers.
The *achieved surface* data are representing, in the case of CIRCOM, the surface compacted by the machine, to which is associated the speed and the indication whether the roller is vibrating or not. More attributes about the compaction energy of the machines can also be added: amplitude, frequency, etc.

CIRPAV
CIRPAV [3] is a little more complicated since it provides two main functions: assistance to the driver for the steering of the machine, thanks to a color display similar to the CIRCOM's one, and automatic elevation control of the screed. The positioning component, called CIRPAV POS, can be based upon different technologies, but has to provide both position and attitude of the screed. The preparation of the mission and the transmission of the design data are still made on the CIRC GS and the on-board processing is done by the CIRPAV OB.
The CIRPAV data flow analysis is presented on Figure 3.

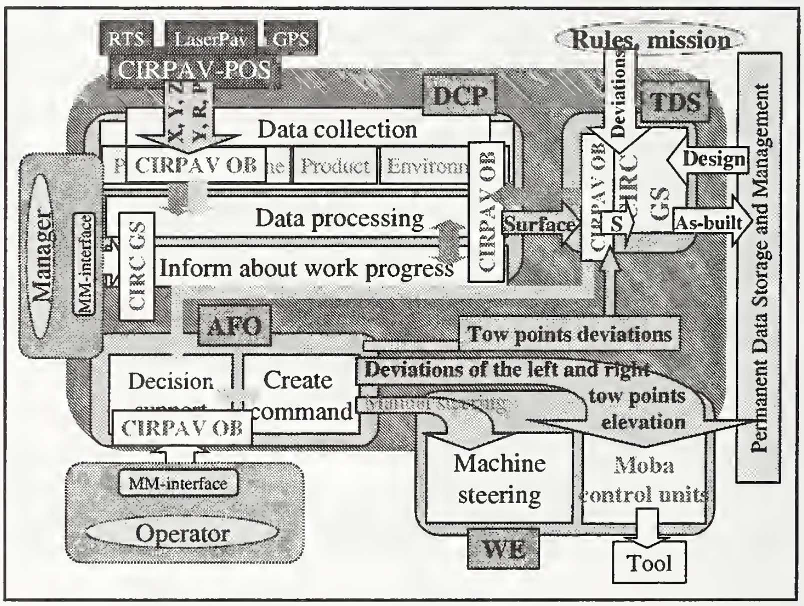


Figure 3: The CIRPAV data flows

Most of the data flows are similar to the CIRCOM's ones, albeit with a different content: *design*, *as-built*, *mission*, *position*, and *achieved surface* data.
The main difference is in a new data flow that is exchanged between AFO and WE (which exists this time due to the automatic control ensured by the CIRPAV SIS): the deviations between the actual tool position and the target tool position.
This data flow category we can call: *deviations for machine control.*

OSYRIS
OSYRIS [2] is an even more complex system where more data flows are addressed, as it includes work execution functions and quality documentation functions. It is a system also where there are several data exchanges between different kind of machines (paver to rollers).
OSYRIS system is mainly designed for standard asphalt pavement sites, executed with 1 paver and 2 compactors. A basic configuration is proposed, with additional

options for on-board functions and office functions.
- The TDS sub-system functions are executed by several pieces of software that are physically located in office computers or in on-board computers, depending of the version of the product.
In the 'full' version of OSYRIS, the whole package includes the office *Design and Documentation* piece of software that allow a full mission description and a full documentation of the achieved work consistent with the geographical data base (OSYRIS Product Model).
These pieces of software can also be located at the level of the *Permanent Data Storage and Management* system which is outside of our SIS, since they are also designed to manage the permanent information after the execution of the work.
- The DCP functions are executed mainly by the on-board computers, located on the paver (*Paver OB*) and on the rollers (*Roller OB*). This means that wireless exchanges are addressed, for instance for temperature data, position data, etc. The visualization of the progress of the work is ensured by the office *Work Documentation* software component (WD) and by the *Work Site Web* (WW) service.
Many on-board sensors are also concerned in OSYRIS, as inputs to the DCP and these data flows are necessary to be standardized .
As far as AFO and WE are concerned, the functions are quite closed from those of CIRC system.
The OSYRIS data flow analysis is presented on Figure 4.

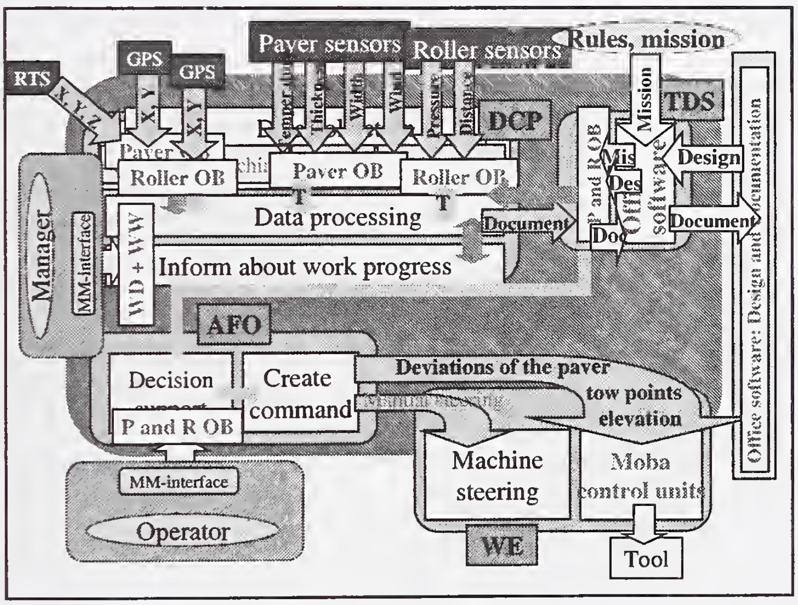


Figure 4: The OSYRIS data flows

The new data flows that are involved in this kind of SIS are material-related measurements that are carried out on the machines, for documentation, operation or transmission to other machines purposes, like: temperature, thickness, width and volume of the layer, environment parameters.
Let us call this new category: *machine sensors data.*

5 SYNTHESIS ABOUT THE DATA FLOWS TO STANDARDIZE

Our analysis, from road construction SIS, has yielded the following 6 main data flows categories:

1. *Design data,*
2. *Mission data,*
3. *Real-time Position and speed data,*
4. *Real-time machine sensors (and achieved surface) data,*
5. *Real-time deviations for machine controls,*
6. *As-built and work documentation data.*

(As *achieved surface data* are obtained from sensors on-board of the machine, they can be grouped with *machine sensors data. Position and speed data* are for the moment considered apart, given their very specific importance).

These categories have different time scales:
- 1 and 6 have very long evolution periods (from several days to several months),
- 2 has medium long evolution periods (from several hours to several days),
- 3, 4 and 5 have short evolution periods (from some milliseconds to several seconds).

Type 6 is the counter part of types 1 and 2, *as-built* corresponding to *design* and *work documentation* to *mission*. These 3 categories are composed of static data. As far as format is concerned, XML seems to be an excellent candidate given its universality and flexibility. In this respect, the new LandXML standard under development [5], proposing both data model and data format should be considered, although non addressing documentation data so far. Raster formats are also to be considered for presenting synthetic documentation maps.

Categories 3 to 5 are real-time data, attached to the machine itself. An important remark is that these data have necessarily to be processed with respect to the position of the machine (this is obvious for category 3). An interesting way of representing these data is to use the *Ribbon* data base structure which has already been used in both CIRC and OSYRIS projects and is described in [6] and [7]. Ribbons allow the description of geometry, this way they can be used also for design and as-built data, but they can also store parameters with respect to the position of the machine. These parameters can be either machine sensors data or deviations for tool control. Ribbons are generated by a generator moving along an oriented polyline and allow parameter interpolation as wall as curvilinear transformation. They are time-tagged and each machine is generating its own ribbon. All the different ribbons can be merged to output the synthetic maps for as-built documentation.

6 CONCLUSION AND PERSPECTIVES

Our analysis has brought to the fore several data flows categories that must be considered separately since they play different roles in the process. Still these different categories need to be described more precisely to continue the work, but experience from European projects such as CIRC or OSYRIS already tells us that the *Ribbon* structure can be considered as a universal digital model for real-time data. As far as static pre-execution and post-execution data are concerned, existing standards as LandXML or raster format must be used.

7 REFERENCES

[1] Peyret F. Miyatake H., "Standardization of information systems for civil engineering sites", *Proc. 18th International Symposium on Automation and Robotics for Construction*, Krakow (2001), 235-240

[2] Ligier A., Fliedner J., Kajanen J., Peyret F., "Open system for road information support", *Proc. 18th International Symposium on Automation and Robotics for Construction*, Krakow (2001), 117-182

[3] Peyret F. , Jurasz J., Carrel A., Zekri E., Gorham, B., "The Computer Integrated Road Construction project", *Int. journal Automation in Construction* , vol. 9 (2000), 447-461

[4] Pampagnin L.H., Peyret F., Garcia G., "Architecture of a GPS-based guiding system for road compaction", *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Liège (1998)

[5] http://www.landxml.org/

[6] Jurasz J., "Universal digital environment model for computer integrated road construction", *Proc. 18th International Symposium on Automation and Robotics for Construction*, Krakow (2001), 37-42

[7] Jurasz J., Ligier, A., Horn A.., "On-board data management structure for advanced construction machine support", to be published in *Proc. 19th International Symposium on Automation and Robotics for Construction*, Washington DC (2002)

PARTS AND PACKETS UNIFICATION FOR CONSTRUCTION AUTOMATION AND ROBOTS

Tomohiro UMETANI*, Tatsuo ARAI*, Yasushi MAE*, Kenji INOUE*, and Jun-ichiro MAEDA**
*Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka, Osaka, 560-8531, JAPAN
**Shimizu Co.
umetani@arai-lab.sys.es.osaka-u.ac.jp
+81-6-6850-6368, +81-6-6850-6341(FAX)

ABSTRACT: This paper proposes a method of configuration of database system for parts oriented construction. In parts oriented construction, robots acquire information related to the tasks from the server through ID attached to the construction components. This system aims (1) simplification in motion planning of the construction robots using information attached to components, (2) reconfiguration of data structure for process of construction tasks. The robot performs the tasks according to local planning based on the local information. The operation server reconfigure parts information based on task result by robots so that the next robot can obtain the required information from the composed components in the next phase. We call the parts integrated information "packets", and call our idea "parts and packets unification." The feasibility of our proposed method is also shown in the preliminary experiments where an RFID tag attached to a component is applied for data exchange.

KEY WORDS: Automatic Construction, Information Integrated Components, Data Collection System and Labeling, Automatic Reconfiguration of Parts Information, RFID Tags

1. INTRODUCTION

The recent advancement of information and communication technologies has brought feasibility of efficient construction automation [1][2][3]. However hard problems of construction automation still exist. In construction site, workspace changes according to a progress of construction works and works are performed in parallel. And a building is one-article production. Therefore, it is difficult to make a global and perfect plan of the construction robots beforehand.

It is reasonable that the parts-oriented construction, where a robot makes a plan using information of parts on site, is applied to the construction automation. For the parts-oriented construction, the server manages a relation between construction components and their information. And the server has to manage the information of working robots. By managing the relationship between the parts and their information, the robot can check the assembled parts and can correct errors caused by disturbances. Therefore, it is expected that the work may not be repeated to cancel the errors and the efficiency may be improved of the work.

In the construction tasks, new components are generated in the progress of construction. We call this newly composed component a "module," since it is used as a part of the building. A module is used as a component in the next phase. Attributes of a module (form, weight, center of gravity, etc.) will change as components are composed to a new module. It is difficult that a robot obtains required information of each component composed in a module. Then parts information about the module should be reconfigured. In order to reconfigure the parts information automatically, it is necessary to match a component (parts) with its accompanying information (packets). We call this state "parts and packets unified state." We propose a parts oriented construc-

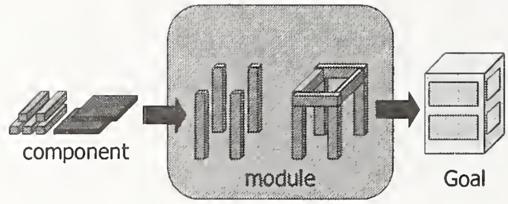


Fig. 1 Change of components in construction

tion system using the concept of “parts and packets unification.”

In this study we consider parts and packets unification, where every object, part, module, and structure carries its related information on itself with data carrier such as RFID tags [4][5], bar codes [6], and so on. We are also interested in its effective usage of the construction automation and robots. In the following we will discuss what a feasible database will be, how attributes will be dealt with as construction objects are processed, and what physical and information interactions between objects and robots will be, by citing some example tasks.

2. PARTS ORIENTED CONSTRUCTION

2.1. Unification of Parts and Their Information

We show a construction process of a building in **Fig. 1**. In the construction tasks, information and attributes of each component and those of the completed building are known. And the work procedure is also known. A new module is generated by composing the components. Attributes of a module (form, weight, center of gravity, etc.) will change as the module is composed. But their information is unknown. In a process of construction, attributes of a module should be clear for local planning of the construction robots. We propose a method to clarify the attribute of the module using an ID attached to each component.

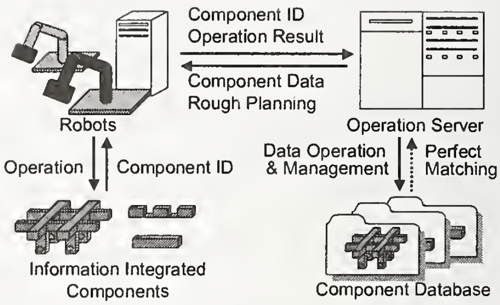


Fig. 2 Diagram of parts oriented construction

2.2. Parts Oriented Construction

We consider a method of realization of the parts oriented construction system using information-integrated components. In **Fig. 2**, we show a relation between construction robots, components, and their information. A Robot acquires the parts information via ID attached to the components, and operates the components. The robot sends the components ID and the task results to the operating server while achieving the tasks. The server sends the parts information to the robot by the components ID. There are several methods to obtain the components ID in this system, for example, RFID tags, bar code, and so on. And there are several ways to communicate between the robots and the server, for example, a private network, LAN, and so on.

A method of informational operation by the server and acquisition of the parts information by the robots are serious problems to the settle of the construction system. Thus we discuss the method of the informational operation and management of the parts information in section 3. And in section 4, we carry out experiments of the acquisition of the parts information using RFID tags. We discuss the problem in the parts oriented construction system.

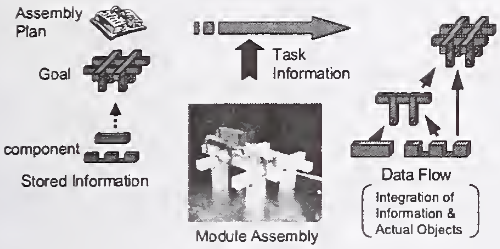


Fig. 3 Example of parts and packets unification

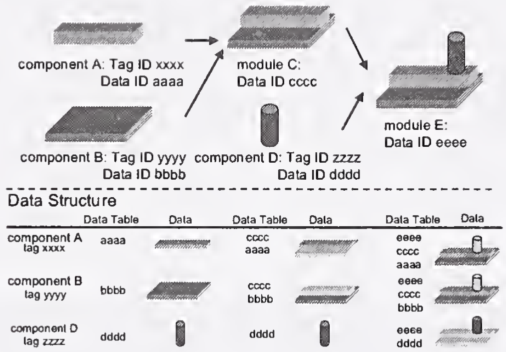


Fig. 4 Autonomous data renewal process in module assembly

3. MANAGEMENT OF STRUCTURE OF PARTS INFORMATION

3.1. Autonomous Data Acquisition

Fig. 3 shows a concept of autonomous composition of parts information. Robots achieve a construction task based on the task plan. The parts information is also composed from the task results. Thus the parts information of the modules is also composed. Fig. 3 shows an example of how a module and its attributes are reconfigured through an assembly task. The attributes of a real module and those stored as parts information of the module in the server are matched by autonomous reconfiguration of the parts information. Thus, the robot can determine the detail motion plan using the rough motion plan and the stored information of the modules.

We show the merits of the autonomous data reconfiguration. If the attributes of modules are stored as the accompanied parts information, the robot can obtain the information of the modules easily by local information acquisition through the ID.

3.2. Properties of Parts Information

Attributes of a module, for example, weight, center of the gravity, posture, and so on, change in robots and machine acting on a module. It is expected that recognition of the stored attribute utilizing the task result of robots lead to consumption of lower energy and safety of the construction site for workers, machine, and robots.

In this section, we describe the property of the attributes of modules. The attributes of modules change in a progress of tasks. Some attributes inherit the former attributes before the change (dependent), and others do not inherit the former attributes (independent). For example, form, moment, and the properties of a module depend on those of a former module. On the other hand, appearance and accuracy of a module do not depend on those of a former module. Accuracy of attachment and processing of modules are different since abilities of workers or robots are different. And new attributes are added when a module is processed, i.e. welding and cutting.

The attributes of each module should be reconfigured during construction. A large module consists of many modules.

3.3. Reconfiguration of Parts Information

Fig. 4 shows the change of the parts information stored in the database in the process of module assembly. The top of the figure shows an example of the process and how each component is composed is shown. The bottom of the figure shows the parts information of the composed module. In this figure, module C is composed from component A and B. And module E is composed from module C and element D. The code "xxxx" indicates the ID code of each tag or information. In this example, the module is composed so that the parts information is also reconfigured. It is desired that the database should be set up in order that the robot can refer the same information of the module via each tag attached to the module. We will discuss the structure of the parts information and the method of reconfiguration of reference of the parts information.

We show the method of the configuration of the parts information in the following. It is supposed that a tag is attached to component A, B, and D. The parts information referred to each tag is stored in the server. When the robot composes component A and B, the robot acquires the parts information using the reference from the tag xxxx attached to component A and tag yyyy attached to component B, and the robot achieves the task. Then both attributes of component A and B change so that the newly reconfigured information of module C is generated. And the attribute of module C is referred to the tags attached to component A and B. Therefore the robot can acquire the information of the module via any tags attached.

Fig. 5 shows the reconfiguration of data reference from tag xxxx, yyyy in the composition process. The left side of Fig. 5 shows the ranking of the parts information references. For the first time, the reference of parts information about each element is stored in the reference map. In the process of the module composition, the ranking of the reference changes in order that the robot refers to the information of the composed module via each tag attached to the module. Through the data operation, the robot can refer to the present information of the module via any tags attached to the module. On the other hand, the robot reaches the information of each element using the list of the reference map for each tag.

We make a simple program of reconfiguration of the database, so that the robot can reconfigure the parts information of the module.

4. EXPERIMENT OF ACQUISITION OF PARTS INFORMATION

4.1. Acquisition of Parts Information by Robot

We describe a method of acquiring parts information by a robot. To obtain the information attached to the parts, a robot recognizes letters or markers printed on the parts in a method [7]. But a vision sensing has a difficulty to acquire the information, since a vision sensor is strongly influenced by conditions, the surface status of parts and a light source.

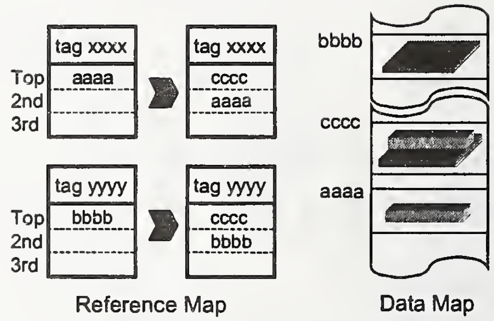


Fig. 5 Reconfiguration of data reference in module assembly

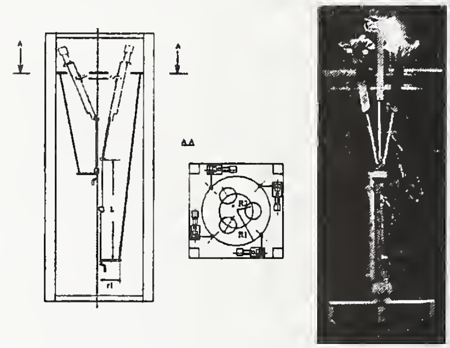


Fig. 6 Prototype of hybrid arm

We adopt an RFID tag as a medium for acquisition of the information. An RFID tag is robust against a change of environmental conditions, e.g. temperature and light sources. However it is influenced by a metal that is frequently used as parts in construction site, since they are communicated by radio signal. Here we try to carry out experiments of acquiring parts information using RFID tags to ascertain the feasibility. Through the experiment, we discuss possibility of autonomous construction using RFID tags.

The experiment setup includes using an antenna, RFID tags, parts models, and the prototype hybrid parallel arm for autonomous construction (see **Fig. 6**). We consider the conditions about a robot that acquire parts in-

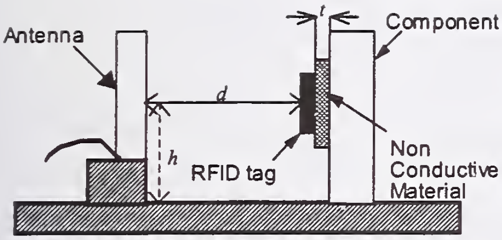


Fig. 7 Experimental environment

formation from the parts by RFID tags, and also discuss how the tags should be attached.

4.2. Experimental Setup

The experimental setup is shown in **Fig. 7**. The center of the antenna is set at the same level, h, as the RFID tag, and the distance between the antenna and the RFID tag is *d*. and the distance between the RFID tag and the component is *t*. The RFID tag is OMRON V700-D13P21. The antenna is OMRON V700-H01. In this setup, the height *h* is 215[mm]. The component is 60[mm] by 60[mm] by 130[mm], and it is made of iron. There are two pins on a pair of the parallel surface. The length of each pin is 15[mm]. The floor is made of metal.

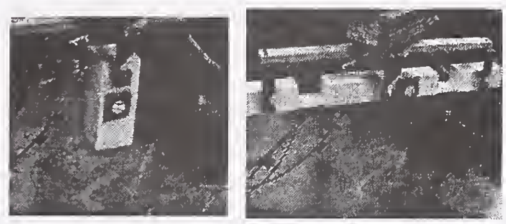
We examine what the maximum distance between the antenna and the RFID tag attached to the component where no antenna can communicate with the tag. And we measure the thickness between the RFID tag and the component to obtain the reasonable communicable distance. A nonconductive material is introduced to cancel the metal influence on radio signal. Its thickness will be varied. To compare the ability of the communicable distance, we measure the communicable distance between the antenna and the RFID tag attached to the nonconductive object.

4.3. Measurement Results

We measure the communicable distance by changes of the distance *d* between the RFID tag and the antenna. We wrote small-size data to the RFID tag, and read the data from the tag using the antenna. We regard that the trial succeeds if both writing and reading tests are succeeded. For each distance *d*, 10 trials are carried out, and we set distance *d* by 5[mm]. We regard the distance that all trial at the distance succeeds as the communicable distance. The thickness of the nonconductive material *t* is set 0, 3, 6, and 9[mm].

Table 1 Communicable distance between antenna an RFID tags

t [mm]	0	3	6	9
Maximum Distance [mm]	Failure	90	125	130



(a) Pillar (b) Beam

Fig. 8 Information integrated components handled by the hybrid arm

The communicable distance between the antenna and the RFID tag is shown in **Table 1**. The table shows, a reasonable communicable distance can be defined if the RFID tag is attached to the component with the nonconductive material, but not directly. In the comparison, the communicable distance in the case the tag attached to the nonconductive material is 180[mm].

4.4. Acquisition of Parts Information in Manipulation

We examine acquisition of parts information while the hybrid arm shown in **Fig. 8**. The hybrid arm manipulates the parts and the experimental setups are carried out in the same conditions as in section 4.2. The height *h* from the ground to an RFID tag is set 190[mm] so that the hook or the end-effector of the hybrid arm might not become obstructive on manipulating a component. We attach an RFID tag to a beam component at 40[mm] shifted from the center since the end-effector has the hook at the center. In manipulating the beam component, the height h is set 230[mm]. We set the thickness between the component and the tag t is set 6[mm] from the measurement results in section 4.3. Fig. 8 shows how the arm can handle the component.

We tried the communication test with the distance 80[mm]. From this examination, the robot can acquire the parts information via the RFID tag attached to the component using the antenna. But, when we shift more 10[mm] from the center of the arm or set the communicable distance d more 20[mm], the antenna cannot communicate with the RFID tag. Therefore, the distance should be carefully selected in the actual application.

5. CONCLUSIONS

This paper proposes a parts oriented construction system using the concept of "parts and packets unification." We propose an autonomous data reconfiguration using the change of attributes of modules that a robot acts. We discussed the data structure of the parts database, reconfiguration of the parts information changes on task procedures, and the property of parts information. And we showed the autonomous data acquisition using RFID tags attached to components. Through the experiment, we assured the feasibility of the proposed parts oriented construction system.

ACKNOWLEDGMENT

This work has been done in IMS international joint research program IF7-II. We appreciate the useful discussion with Mr. Jun-ichi Yagi, Shimizu Co. and other members of the program.

REFERENCE

[1] T. Arai, K. Yuasa, *et al.*: "A Hybrid Drive Parallel Arm for Heavy Material Handling," IEEE Robotics and Automation Magazine, Vol. 9, No. 1, pp. 45 – 54, 2002.

[2] "Study on innovative and intelligent field factory domestic feasibility study for IMS," IMS Promotion Center, 1998.

[3] H. Osumi and T. Arai: "Cooperative Control between Two Position-controlled Manipulators," Proceedings of the IEEE International Conference on Robotics and Automation, Vol. 3, pp. 1509 – 1514, 1994.

[4] K. M. Furlani and W. C. Stone: "Architecture for Discrete Construction Component Tracking," Proceedings of the 16th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction, pp. 289 – 294, 1999.

[5] K. M. Furlani and L. E. Pfeffer, "Automated Tracking of Structural Steel Members at the Construction Site," Proceedings of the 17th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction, pp. 1201 – 1206, 2000

[6] J. Ota, M. Yamamoto, *et al.*: "Environmental Support Method for Mobile Robots Using Visual Marks with Memory Storage," Proceedings of the 1999 IEEE International Conference on Robotics and Automation, pp. 2976 – 2981, 1999.

[7] K. Sawa, W. Ohyama, *et al.*: "Recognition of Degraded Character String on Steel Plate," The Transactions of the IEE Japan, Vol. 122-C, No. 1, pp. 84 – 89, 2002, (in Japanese).

An UML-XML-RDB Model Mapping Solution for Facilitating Information Standardization and Sharing in Construction Industry

I-Chen Wu[1] and Shang-Hsien Hsieh[2]
Department of Civil Engineering, National Taiwan University
Taipei 10617, Taiwan, R.O.C.

kwu@caece.net[1] and shhsieh@ce.ntu.edu.tw[2]

Abstract: To facilitate information standardization and sharing in Construction Industry, this paper presents a simple but effective approach that maps the UML (Unified Modeling Language) object-oriented information model related to a construction project to an XML schema, then to a Relational DataBase (RDB) schema. First of all, the mapping between UML model and XML schema is discussed since UML has been a popular tool to model the static structure and dynamic behaviors of the information and processes in a construction project, while XML has become a de-facto standard for information sharing and exchange. Then, a set of consistent rules for mapping from XML schema to RDB's Entity-Relational (E-R) model are studied and established since RDB has been the most popular choice for information management. The present study focuses on making the set of rules simple and easy-to-implement for most applications in construction industry. Finally, a mapping tool for automatically generating RDB schemas from XML Schemas is developed.

Keywords: UML, XML Schema, RDB, Information Standardization and Sharing

Introduction

During the life cycle of a construction project, voluminous data and information are usually created along the delivery processes of construction products. There is always a need to share and exchange these engineering data and information among related parties involved in the project. However, because data models defined in the information management systems of various parties are usually different, it is always difficult to directly map from one data model to another for the purpose of information sharing and exchange.

To address the aforementioned information-sharing problem among various parties, standardization of the data model for a construction project is usually inevitable. Due to the popularity of object-oriented modeling approach in recent years, an object-oriented information model is often constructed to represent the static structure and dynamic behaviors of the information and processes in a construction project and expressed by UML (Unified Modeling Language) [1], a popular tool for object-oriented modeling. Moreover, Extensible Markup Language (XML) [2] has become a de-facto standard for information sharing and exchange in recent years. Therefore, there is a need to define an XML schema based upon the UML object-oriented information model to further facilitate information sharing,

Furthermore, due to the popular use of Relational Database (RDB) technique for information management, it is also important to address the mapping between an XML schema and an RDB model. However, the mapping is not an easy one because the data model of an XML document is fundamentally different from that of a relational database. Especially the structure of an XML document is hierarchical and the XML elements may be nested and repeated.

Although several RDB providers have provided common data import/export tools to allow for data transformation between the XML documents and the RDB tables, the capabilities of these tools are still quite limited. Recently some commercial software has also provided tools for transforming information in XML documents into RDB. However, most of them can only transform simple XML documents. That is, if the XML documents have a nested data structure and association, most of these tools are still incapable of making a complete transformation.

In addition, several XML-to-Relational transformation algorithms and mapping tools have been proposed in the literature. For example, Bourret et al. [3] introduced an XML-RDB mapping language to specify transformation rules for generating an RDB schema from an existing XML DTD (Document Type Definition). They also proposed a lightweight, DBMS- and platform-independent load/extract utility to facilitate the data transfer between XML documents and relational databases. Lee and Chu [4] proposes a method where the hidden semantic constraints in DTD are systematically found and translated into relational data models. However, most of them deal with XML DTD, instead of XML Schema, which is more flexible and offers more supports for data types.

This paper presents an UML-XML-RDB model mapping solution that maps the UML object-oriented information model to an XML schema, then to an RDB schema. First of all, the mapping between UML model and XML schema is discussed. Then, a set of consistent rules for mapping from XML schema to RDB's Entity-Relational (E-R) model is presented. In this work, the set of rules is made simple and easy-to-implement for most applications in construction industry. Finally, a mapping tool developed for automatically generating RDB schemas from XML Schemas is discussed.

Constructing XML Schema from UML Data Model

This work adopts the concept of the XML Metadata Interchange specification (XMI), which defines a rigorous approach for generating an XML DTD from a metamodel definition, and slightly extends the approach of XMI for mapping object-oriented data model expressed by UML to XML Schema. The transformation rules employed in the mapping process are discussed as follows:

1. Mapping UML Classes to XML Elements

The UML Classes show the structural and behavioral features in the object-oriented Model. These features include attributes, association, aggregation, and composition. On the other hand, XML elements serve as a container for attribute and child elements. Thus, mapping UML classes to XML elements are quite straightforward.

2. Mapping UML Attributes to XML Attributes or Elements

Basically, either a primitive data type or an enumeration of UML attributes may be represented as an XML attribute. However, XML parser removes all extra whitespace characters, such as tabs, linefeeds, etc. That makes XML attributes mainly appropriate for simple datatypes of short string values. On the other hand, one can map attributes of an UML class to separate child elements of the corresponding XML element of the class.

3. Flagging UML Object Relationships by an XML Attribute

The current version (Version 1.0) of XML Schema does not yet have direct and full supports for expressing a complete object-oriented model, especially the distinction between the delegation and aggregation relationships that commonly exist in the model. Therefore, this work employs a special attribute named "relation" with a value of either "delegation" or "aggregation" to flag the relationship between the owning XML element and its child elements. It will be shown later in this paper how this special attribute is used to help mapping the XML object schema to a RDB table schema.

4. Constraints on Naming XML Elements

In general, the UML class name is directly used as the XML tag name in the mapping process. However, there are certain constraints we must comply when naming the XML elements:

- The tag name cannot have spaces in it, but symbols like ".", "-", and "_" are allowed.
- The tag name should not start with the string "XML".

Mapping XML Objects to Relational Database Tables

XML provides many of the things found in databases, such as storage (e.g., XML documents), schemas (e.g., DTDs, and XML schemas), query languages (e.g., XQuery, XPath, XQL), programming interfaces (e.g., SAX, DOM, JDOM), and more. However, XML lacks many important features that are supported by the real databases, such as efficient storage, indexes, security, transactions and data integrity, multi-user access, triggers, queries across multiple documents, etc. [5] Therefore, databases are still the top choice for managing information in a enterprise or government agency. Although several kinds of database technologies are currently available in the market (e.g., object-oriented database, object-relational database, etc.), RDB technology remains the most popular one employed in the construction industry.

On the other hand, there is a trend for using XML documents with a standardized schema as a medium for information sharing and exchange among different parties involved in a construction project. In order to transfer data from XML documents to an RDB for efficient management, it is necessary to map the XML document schema (XML Schema) to the RDB table schema.

Because the XML Schema can be viewed as a tree of objects that is mapped from the object model of its corresponding UML class diagrams (as already discussed in the previous section), this work employs the following object-relational mapping rules to map the XML schema to RDB tables [6]:

1.Mapping Elements to Tables

An XML Element object with attributes, element content, or mixed content (i.e., a complex Element object) is mapped to a separate RDB table as shown in Fig. 1.

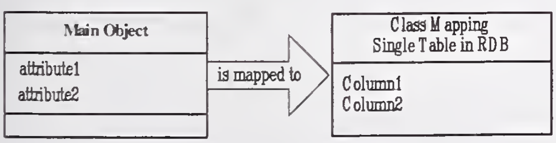


Fig. 1 Mapping an XML Element to an RDB table

2.Mapping Attributes to Columns

Since an XML Element is mapped to an RDB table, the attributes of the XML Element is naturally mapped to the columns of the corresponding RDB table automatically (as shown in Fig. 2).

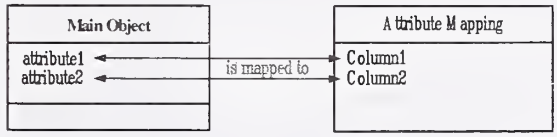


Fig. 2 Mapping attributes of an XML Element to the columns of its corresponding RDB table

3.Mapping XML Elements with Delegation or Aggregation Relationship to RDB tables

For mapping the delegation (or dependency) relationship between XML Element objects, the parent object and its child objects are mapped to separate RDB tables with a primary key and a foreign key, respectively, for associating them (as shown in Fig. 3).

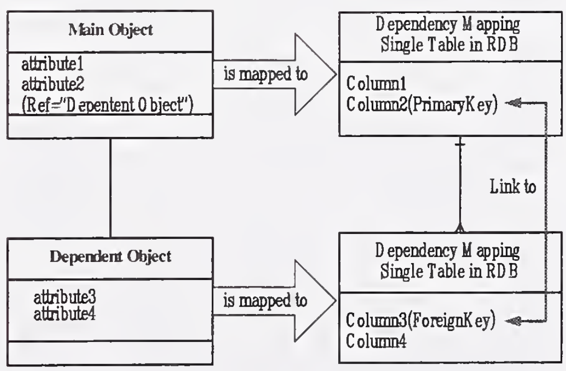


Fig. 3 Mapping XML Element objects with delegation (or dependency) relation-ship to RDB tables

For mapping the aggregation relationship between XML Element objects, the parent object and its child objects are mapped to a single RDB table uniting attributes of all the XML Element objects (as shown in Fig. 4).

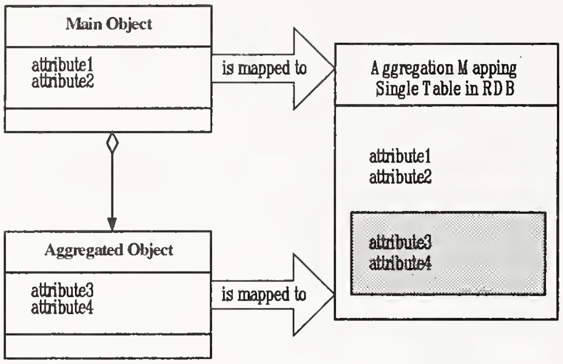


Fig. 4 Mapping XML Element objects with aggregation relationship to a RDB table

XML-RDB Mapping Tool

In this work, an XML-RDB mapping tool, called Schema2RDB, has been developed to verify the mapping rules discussed in the previous section and to automate and ease the task of creating RDB tables based on an existing XML schema. The tool is implemented using Java solutions mainly due to its platform-independent capabilities. JDOM API [7] is employed for an easy and efficient reading, manipulation, and writing of an XML document. JDOM itself is not a parser; instead, it is a wrapper. Therefore, it requires the presence of an underlying parser. In this work, Xerces is used for the parser. Because JDOM is used to handle all types of XML Data in this work, we can say that Schema2RDB is a 100% pure Java application.

Figure 5 shows the interactions among the user, Schema2RDB, and an RDB. First of all, the user assigns an XML Schema file and executes Schema2RDB. Then, Schema2RDB starts parsing the XML Schema file, applies the mapping rules discussed in the previous section, and generates SQL statements automatically via JDBC [8] to create relational database tables (in this work, Microsoft SQL Server). Finally, the RDB responds with the information about the created database tables and Schema2RDB forwards the information to the user.

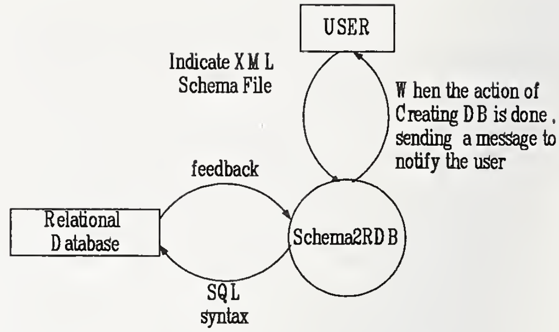


Fig. 5 Interactions among the user, Schema2RDB, and an RDB

The process of mapping an XML schema to RDB table schemas in Schema2RDB is discussed in more detail as follows:

1. As shown in Fig. 6, the structure of the XML schema is first parsed by a XML parser and converted into an XML DOM. The parser provides a way to extract XML schema information and can automatically resolve names of elements and attributes.

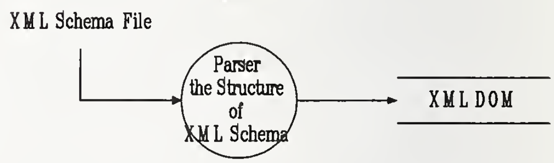


Fig. 6. Parsing the structure of XML schema

2. After the parser has resolved the names of the elements and attributes, Schema2RDB determines the names of the tables and columns as well as relationships for mapping to a relational schema.

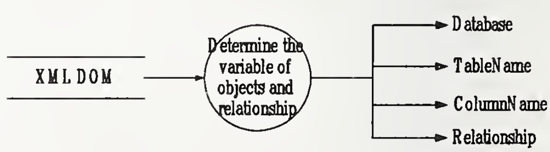


Fig. 7 Determination of the names of the tables and columns as well as relationships for RDB

3. Schema2RDB then composes the SQL statements required to create the RDB tables corresponding to the parsed XML schema (as shown in Fig. 8). The template of a SQL statement is first constructed according to the variable, then the value of

the variable is used to complete the composition process.

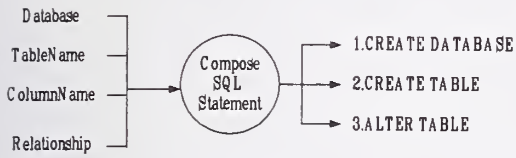


Fig. 8 Composing a SQL statement

4. Through the JDBC driver, Schema2RDB creates a connection to the relational database, executes the SQL statements, and finally closes the database connection. At this point, database tables are ready in the RDB for accepting data from the corresponding XML documents.

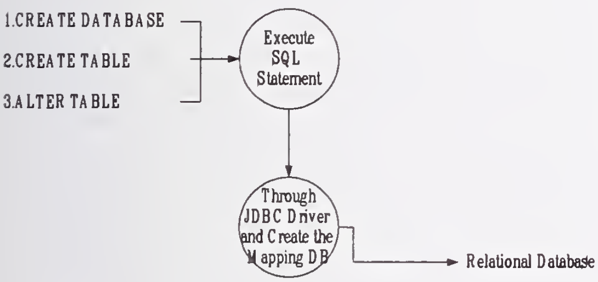


Fig. 9 Process for creating RDB tables

In addition, Schema2RDB have been successfully tested by several verification examples. Although these examples are not considered sophisticated ones, the effectiveness of Schema2RDB in mapping XML Schema to RDB table schema has been observed.

Conclusions

This paper has presented a simple solution for mapping UML object-oriented model to XML Schema, and then to RDB table schema. The framework and process of the mapping approach have been discussed herein. A set of simple, consistent, and easy-to-implement mapping rules has been proposed. In addition, a mapping tool, called Schema2RDB, has been developed to facilitate automatic generation of RDB table schemas from XML Schemas. It is hope that the work presented in this paper can help facilitate information standardization and sharing in construction industry

REFERENCES

1. M. Fowler and K. Scott, UML Distilled, 2nd Edition, Addison Wesley, Boston, 2000.

2. W. J. Pardi, XML in Action, Microsoft Press, Washington, 1999.

3. R. Bourret, C. Bornhövd, and A. Buchmann, “A Generic Load/Extract Utility for Data Transfer Between XML Documents and Relational Databases,” Proceedings of the Second International Workshop on Advance Issues of E-Commerce and Web-based Information Systems (WECWIS 2000), San Jose, California, U.S.A., June 8-9, 2000.

4. D. Lee and W. W. Chu, “CPI: Constraints-Preserving Inlining Algorithm for Mapping XML DTD to Relational Schema,” Data Knowledge Engineering, Vol. 39, No. 1, 2001, 3-25.

5. R. Bourret, “XML and Databases,” from web page at http://www.rpbourret.com/xml/XMLAndDatabases.htm, February, 2002.

6. I. C. Wu, “Information Sharing and Exchange in Building Administration Using XML Technology,” M.S. Thesis, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan, 2002.

7. J. Britt, “Interactive Java & JDOM online tutorial,” Commercially available from the TopXML website at http://www.topxml.com/, 2002.

8. G. Reese, Database Programming with JDBC and Java, 2/e, O'Reilly, U.S.A., 2001.

Using Customized Navigational Models to Deliver more Efficient Interaction with Mobile Computing Devices on Construction Sites

Jan Reinhardt,
Research Assistant, Carnegie Mellon University, Dept. of Civil and Environmental Engineering,
janr@andrew.cmu.edu

Burcu Akinci,
Assistant Professor of Civil Engineering, Carnegie Mellon University,
bakinci@andrew.cmu.edu

James H. Garrett, Jr.,
Professor of Civil Engineering, Carnegie Mellon University,
garrett@cmu.edu

Abstract
Mobile computing devices have started to become useful tools on construction sites. These devices can make large amounts of data available and support data collection tasks on a construction site. Even though battery life, computing power and data storage capacity have greatly improved and will probably not be limiting factors for mobile computing solutions, it is still difficult to interact with mobile computing devices efficiently. Some of the factors limiting efficient interaction are the size of the screen of a mobile computing device, the deficiency of a keyboard and the limited accuracy of the pointing device. The approach we pursue is to limit the interaction needed for specific tasks that are carried out using a mobile computing device. To do this we are working on developing *Navigational Models*, which reduce navigation through large data sets by grouping data entities closely together, according to the need to be accessed for a specific task. Moreover, *Navigational Models* appear to be a useful construct to enter and access data at the right level of detail, which they needed. In this paper, we discuss the need for more efficient interaction with mobile computing devices on construction sites and introduce *Navigational Models.*

KEYWORDS: Construction, Human Computer Interaction, Mobile Computing, Navigation through Complex Data Structures, Product and Process Models

1 Introduction

Many documents available at a site office, e.g. drawings, specifications and schedules, are also needed on a construction site. Currently, for the "last mile" of the flow of information from the site office to the job site, computer-based documents do not work very well. In a majority of the cases, the information stored within an electronic document, or a product and process model, needs to be printed out to access it on the construction site.

Oftentimes, paper documents used on a construction site are not in a format appropriate for the environment in which they need to be used. Drawings and schedules may be too large to be carried to the locations where they are needed, and bills of quantities can amount to several hundred pages. As a result, site superintendents may not have the data with them at the time when they need the information. This deficiency leads to time-consuming document preparation processes to make data available on a site, and additional data transition processes to reintegrate data that has been acquired on a site into electronic documents.

In recent years, technologies that can bring computing power to a construction site have emerged. Mobile and wearable computers are powerful portable devices that can be used on construction sites. Moreover, Personal Digital Assistants (PDAs) and pocket computers have become popular because of their portability. Information contained in electronic documents becomes accessible right on a construction site when these devices are used. Observations that are made on a construction site can be integrated into electronic documents or models without having to make notes and to enter the information contained in the notes after being back in the office. Current mobile computing solutions for project management tasks in construction projects still have limitations. These solutions limit the complexity of the data with which the user is able to interact efficiently. Unless mobile computing devices for construction sites can bring all information to the construction site – both complex and not so complex information – and provide an environment within which a user can interact with the data efficiently, problematic paper-based solutions will continue to be used as the best available approach.

2 Current solutions

Many vendors of established IT solutions for construction processes, such as Primavera, AutoDesk and Meridian, provide mobile clients for their main product, e.g., Primavera OnTrack, AutoDesk OnSite, Meridian Prolog Pocket, respectively. The client applications are equipped with functionalities needed on a construction site in order to access and collect data for a given application.

The information we acquired from site superintendents, who have been using a commercially available mobile project management solution, suggests it has been hard for them to enter textual data on the construction site in an efficient way. For example, the specification of the location of a problem requires cumbersome selections from long lists of location descriptions. If the list of possible locations was kept shorter they were not able to specify the location at the desired level of detail. A similar selection problem was identified for the input of the problem description. Even though the considered application handled relatively simple data sets, the interaction with the device and the application was challenging, as the site superintendents had to trade off between an accurate description of the problem and efficient data input. Case studies carried out by another research group corroborate our observations that current mobile computing solutions fail to make complex data sets available to the user efficiently (Dogan et al 2001).

3 Related work

There appear to be three areas that are relevant for research addressing the problems stated in the previous section: a) information and data collection needs for knowledge based tasks on construction sites, b) Human Computer Interaction (HCI) methods for interaction with large data sets and mobile computing devices, and c) data models that provide the information needed on the construction site.

Obviously, there are a number of knowledge-based tasks that are needed on construction sites, which may vary among projects, but trades also may vary within one project as the project advances. The published literature does not provide much information that explicitly states knowledge-based processes and data needs on construction sites. Nevertheless, there are a number of implicit statements that information is needed on construction sites and office-based processes require information, which needs to be collected on the construction site (Dogan et al 2001, Pena-Mora, Park 2001; Mills 1999, Tommelein 1998). These statements motivate our research to make the data needed accessible in an efficient way.

For an efficient navigation through large data sets, Furnas (1986) has identified that tree structures are especially beneficial, as they allow the user to interact with the data on different levels of detail and access detail and context information at the same time. Chien and Flemming (2002) propose tree data structures for efficient navigation through large design spaces in architectural design and Bjoerk (1999) states that rich linkage between semantically related entities in a large data set enhances the efficiency of information access (Bjoerk 1999; Chien and

Flemming 2002). These techniques for efficient navigation through large data sets are relevant for our research as the sources of data that may be needed on a construction site can be large (construction projects may have several hundred drawings and thousands of pages of specifications).

Even though product and process modeling originated from design rather than construction problems, researchers have shown the usefulness of product and process models for the construction phase of buildings. There have been proposed product and process models that specifically address processes in the construction phase of a building (e.g., Stumpf 1996, Akinci 2000, Froese et al 1999). These product and process models appear to be the way of handling and storing construction related data in the future and should therefore be the model that contains the data that we want to make available on a construction site.

4 Making data accessible on site

In our research, we hypothesize that allowing interaction with data presented to the user on variable levels of detail (LoD) will enhance the efficiency of user interaction. Moreover, we hypothesize that linking semantically related information and presenting the links to the user minimizes navigation needs during interaction with an integrated product and process model that handles the information needed on a construction site.

4.1 The level of detail problem

In a case study with a mobile computing application, we experienced that the level of detail at which data is presented to the user and in which the user enters data in a mobile computing system has a significant effect on the efficiency of the interaction. This experience is consistent with another case study that investigated the usefulness of an Augmented Reality (AR) System for inspection tasks in Nuclear Power Plants (Dutoit et al. 2001).

For construction projects, many information access and data collection tasks can be seen as problems related to the level of detail. We identified punch list creation, punch list maintenance and construction progress monitoring as knowledge-based tasks that are carried out at construction sites and require interaction with data at different levels of detail. For example, the question about what the punch list items of the project are needed to be resolved in the following two weeks may relate to the overall project, a section of the project, a certain room, a trade or a certain contractor. For progress monitoring, the user needs to enter the observed construction progress in a mobile computing device. The observed construction progress may need to be considered at different levels of detail, such as the building element level, the activity level or a higher-level grouping of activities.

Interaction with the computing device becomes inefficient or ineffective if there is a mismatch between the level of detail, in which the user is interested or the user wants to enter data, and the level of detail provided by the application. For the progress monitoring application we tested, (Reinhardt 2000), the level of detail for the construction progress was the building element level (e.g, "installation of wall 3 on the 3rd floor is 50% complete"). This high level of detail was appropriate for tracking closely the work of subcontractors, who were constantly late, but for other less problematic activities, the level of detail was too high. The site superintendent had to enter the construction progress of all building elements of one activity even though the actual information she wanted to enter was that all other activities of a certain trade were complete (lower level of detail). The number of clicks on the touch screen of a mobile computer to enter the observed construction progress at the building element level rather than, for instance, at the activity level was typically around 3*n times higher, where n is the number of building elements associated with an activity (it takes 3 clicks to update an activity that is assigned to a building element). In the case study, n was typically in a range of 3 to 4. For progress monitoring, we could show that the level of detail in which the user interacts with a mobile computing application has a significant effect on the cycles of interaction needed for a particular task.

4.2 Minimizing navigation needs

Considering the objective to make all information needed on a construction site available through a mobile computing device, it is obvious that large amounts of data are involved. We assume product and process models to be the knowledge base for construction projects in the future. Even though these models integrate information and hence make the contained information available to the user in a unified way, the navigation across data domains can become very extensive. For this reason, efficient navigation through the data is crucial for efficient interaction with the mobile computing application. In order to reduce the cycles of interaction needed for a particular navigation task, it appears desirable to link information items that are semantically related. Product and process models may already have a rich set of semantic links between the information items included in the model. In the personal organizer application PowerView, researchers have linked semantically related items and presented these links as context information to the user (Bjoerk et al 2001). For our research, we want to build on Bjoerk's approach of linking semantically related objects together and investigate its usefulness for the computer supported knowledge-based tasks on construction sites.

5 Example envisioned interaction style

The task is to determine the degree of completion of the installation of windows at the building element level. When the last update of the construction progress was done, the windows on the east side were completed 30% and no windows were installed on the other sides of the building. The current situation on the project for this illustrative example is assumed to be as follows: on the south side of the building, three windows have been installed, whereas the east side is already completed; the work on the two other sides has not started yet.

In order to record the current progress situation, the user needs to capture the changes in construction progress since the last site visit. The product model is organized in a given decomposition hierarchy as it is depicted in Figure 1.

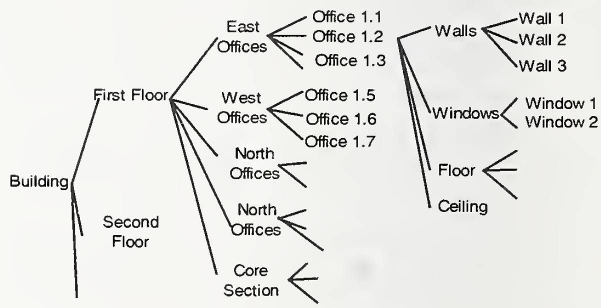


Figure 1, Representation of the data contained in a product and process model

As the user performs update operations only on windows, he/she specifies that a tree structure needs to be derived that only contains information related to windows (see Figure 2). The user also specifies that he/she wants to inspect the windows from outside the building and hence needs, for instance, a tree structure that groups windows at a higher level by the direction of the façade.

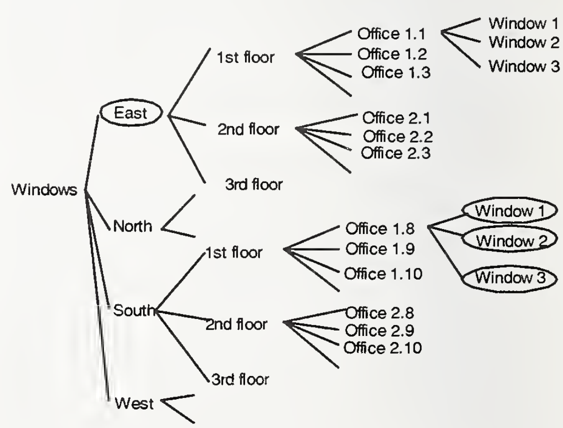


Figure 2, *Navigational Model*, Nodes (in circles) that need to be selected for the progress update

In order to record the progress update for the window installation, the user would select the node **East** (higher level) and classify all building elements on the east-facade as installed. For the south façade the user needs to select the items **Window1**, **Window2**, **Window3** (lower level) and classify the windows as complete. For the south façade, the information that needed to be entered was more detailed and hence the interaction with the system was more detailed compared to the interaction needed to enter the information related to the east façade. The proposed interaction style allows capturing data in a very detailed way, but requires less interaction with the data compared to solutions that present at a static level of detail.

6 Customized *Navigational Models*

Considering the opportunities discussed in the previous sections, we propose a system that supports efficient navigation through large data sets contained in integrated product and process models, which are presented on mobile computing devices. To achieve this, we propose *Navigational Models,* which are constructs that fulfill the following three conditions: a) establish links between information items in product and process models that need to be accessed on the construction site, b) support interaction with data sets on different levels of detail, and c) do not contain information, but rather contain references to information items in a product and process model. Even though linkages between entities may already exist in product and process models, additional linkages or direct linkages specific for a task may become necessary. One representation of a *Navigational Model* is a tree structure, which has leaf nodes that have references to data in a product and process model. The data at the leaves of the tree structure have to have the highest granularity of information in which a user might possibly be interested. Higher-level tree nodes are dynamic aggregations of lower-level nodes and do not have their own attributes other than references to parent or child nodes. However, all data related to the leaf nodes can be made available in non-leaf nodes due to the child-parent associations between the nodes.

One version of the *Navigational Model* supports interaction with data at different levels of detail, as higher-level tree nodes represent data aggregations of lower level data and hence are less specialized. The user is expected to interact with the tree nodes that represent the desired level of detail. In addition to the tree structure representation of a *Navigational Model*, these models should be editable and changeable. The user of the envisioned system should be able to change the configuration of nodes in the model and thus customize the proximity of nodes in the model. The ability to customize a *Navigational Model* responds to the requirement to link information that is semantically related.

The customization of a *Navigational Model* is a means to reduce the navigation space for a given task. In contrast to views of a model, a *Navigational Model* is a construct that specifically supports a task by providing information needed for the task. Views are rather isolated representations of data contained in a model and may not support tasks, as tasks may require efficient interaction with different views. A *Navigational Model* is a task-centered interaction scheme that provides the user with information in representations that are appropriate for the task and the environment of the user. Different from views *Navigational Models* establish linkages between information items contained on product and process models. *Navigational Models* should enhance both information access and data collection on construction sites.

6 Conclusion and future work

Mobile computing devices can support knowledge-based processes in mobile environments and have started to become useful tools on construction sites. These devices can make possible more efficient and consistent data handling during the construction phase. Computing power and future network connectivity will not be limiting factors to make all data needed on a construction site available on electronic devices. Interaction with mobile computing devices is a major limitation for a wider usage of these devices.

Current interaction styles with mobile computing devices and applications often do not consider domain knowledge in order to enhance the efficiency of the interaction with the device. For the domain of project management during the construction phase, we have shown that incorporation of domain knowledge can enhance the interaction with mobile computing devices. Based on these considerations, we have proposed *Navigational Models* that incorporate domain knowledge of the task a user is doing on a construction site. The proposed *Navigational Models* are envisioned to enhance the efficiency of mobile computing applications in two ways: a) allowing interaction with information contained in a product and process model at different levels of detail; and b) link closely together that information upon which the user is likely to execute operations.

This paper has motivated the need for *Navigational Models* by describing the problems users have with levels of detail and navigations through large data sets. We also described the basic functions of such a model and discussed several ways in which these *Navigational Models* might enhance the efficiency of interaction with mobile computing solutions. Future work will implement and test this concept further.

7 References

Dutoit, A.H., Creighton, O., Klinker, G., Kobylinski, R., Vilsmeier, C., Bruegge, B. (2001), *Architectural issues in mobile augmented reality systems: a prototyping case study*, In The Eighth Asian Pacific Conference on Software Engineering (APSEC'2001), Macau SAR, China, December 4-7, 2001.

Bjoerk, S., Holmquist, L.E., Ljungstrand, P., Redstroem, J. (2001), *PowerView: Structured Access to Integrated Information on Small Screens*

Froese, T., Fischer, M., Grobler, F., Ritzenthaler, J., Yu, K., Sutherland, S., Staub, S., Akinci, B., Akbas, R., Koo, B., Barron, A., and Kunz, J. (1999), "Industry Foundation Classes For Project Management-A Trial Implementation", Electronic Journal of Information Technology in Construction, Vol.4, 1999, pp.17-36. Available online at http://itcon.org/1999/2/

Chien, S.-F. and Flemming, U., *Design space navigation in generative design systems*, Automation in Construction, 11 (2002) pp.1-22.

Dogan, F., Wang, F., Lee, G., Jordan J.D., *Onsite Construction Data Input Device*, project description, available under http://swiki.cc. gatech.edu:8080/cs6750b/58, March 2002.

Rezgui, Y. and Debras, P., *An Integrated Approach for a Model Based Document Production and Management*, Electronic Journal of Information Technology in Construction, V. 1 (1996).

Sunkpho, J., J. H. Garrett, Jr., and A. Smailagic, *Opportunities to use Speech Recognition for Bridge Inspection,* In the Proceedings of the 2000 ASCE Construction Congress, pp. 184-193, Orlando, FL, Feb 20-22, 2000.

Buergy, C., Garrett, J. H. Jr., Klausner, M., Anlauf, J., Nobis, G., *Speech-Controlled Wearable Computers for Automotive Shop Workers*, Proceedings of SAE 2001 World Congress, SAE Technical Paper Series No. 2001-01-0606, March 5-8, 2001, Detroit, Michigan.

Reinhardt, Jan; Garrett, J. H. Jr., Scherer, R.J., *The preliminary design of a wearable computer for supporting Construction Progress Monitoring,* Internationales, Kolloquium über die Anwendung der Informatik und der Mathematik in Architektur und Bauwesen, Weimar, Germany, June 2000.

Akinci, B., Fischer, M., Kunz, M., Levitt, R., *Automated Generation of Work Spaces Required by Construction Activities*, CIFE Working Paper #58, CENTER FOR INTEGRATED FACILITY ENGINEERING, Stanford University, June 2000.

SESSION 7

ADVANCED SENSING AND IMAGING TECHNOLOGIES

Performance of Artificial Intelligence Approach on Bridge Coating Assessment

by

Po-Han Chen[1] and Luh-Maan Chang[2]

ABSTRACT: Digital image processing has been prevalently adopted in different areas. In the construction field, image processing has been used for defect detection on steel bridge painting and underground sewer systems. However, non-uniformly illuminated images always cause recognition problems and affect the accuracy. In order to resolve these problems, the neuro-fuzzy recognition approach (NFRA) was proposed. The NFRA segments an image into three areas based on illumination and conducts area-based thresholding. The neural network is used in this approach for automatic generation of three threshold values, with the three average illumination values of the three areas as the input. The fuzzy adjustment is utilized to smooth and adjust the gray level values of the image pixels along the boundaries. In this paper, the framework of NFRA and the rationale of the fuzzy adjustment will be presented, followed by the comparison of the recognition results using NFRA and the multi-resolution pattern classification (MPC) method. The result shows that the proposed NFRA performs fairly well on recognizing rust images. Finally, the conclusions will be drawn.

KEYWORDS: Fuzzy Adjustment; Multi-Resolution Pattern Classification (MPC); Neural Networks; Neuro-Fuzzy Recognition Approach (NFRA)

1. INTRODUCTION

As computerized technologies were widely utilized, digital image processing was also prevalently adopted in many industries (Abraham et al. 1997; Croall and Mason 1992; AbdelRazig et al. 1999; Chen and Chang 2000). In the construction area, image processing has been used for defect recognition on steel bridge painting and underground sewer systems.

There are a number of advantages when using computerized digital image processing. Computerized digital image processing is able to distinguish millions of shades of colors, which are hard to be distinguished by human eyes, and enables the analysis and comparison of images. In addition, digital image processing can accurately calculate defect percentages. Even though digital image processing has such powerful capabilities, there are still some drawbacks that need to be resolved, especially when image quality is poor. Non-uniform illumination is the most frequent problem that accompanies a poor-quality image. In order to obtain better and reasonable results while conducting digital image processing, methods that can diminish the bad effects from non-uniform illumination are needed.

Intelligent surface coating assessment uses neuro-fuzzy recognition approach (NFRA) to resolve the recognition problems resulting from non-uniform illumination. Because the colors in steel bridge rust images are simple, all rust images are converted to grayscale before further processing without losing much information (Chang 2000). The conversion of color images to grayscale images reasonably simplifies the complexity and expedites the recognition process. Then, the illumination-based image segmentation is conducted. In this segmentation, each of the grayscale images is segmented into three areas in accordance with the illumination values of the pixels in the image. The average

[1] Assistant Professor, Nanyang Technological University, Singapore 639798. Tel: +65 6790 4562. Email: cphchen@ntu.edu.sg

[2] Associate Professor, Purdue University, West Lafayette, Indiana 47907, USA. Email: changlm@ecn.purdue.edu

illumination values of the three areas are calculated afterwards and sent to a pre-trained neural network to obtain three corresponding threshold values, which will be used for illumination-based image thresholding later. In the mean time, a fuzzy adjustment is utilized to smooth the threshold values of the pixels along the boundaries between different areas. Finally, image thresholding is applied based on the obtained threshold values to get a binary image containing only the object pixels and the background pixels (Chen 2001).

In order to test the feasibility of the proposed NFRA, the recognition results using NFRA and the multi-resolution pattern classification (MPC) method were compared (Chen 2001).

2. FUZZY ADJUSTMENT

The fuzzy adjustment is to be applied to the image pixels on both sides of the boundaries between areas. Figure 1 illustrates the schematic representation of the fuzzy adjustment. Two inputs are included in this fuzzy system, the "positive difference" and the "negative difference." The output is the "gray level difference." The output is the "gray level adjustment." A set of nine If-Then rules constitutes the kernel of the fuzzy system, as indicated in Table 1 (Chen 2001).

Both the inputs "positive difference" and "negative difference" have three levels: large, medium, and little. The universe of discourse (the range of input) for both inputs ranges from 0 to 20. Differences (both positive and negative) larger than 20 are counted as 20.

The "gray level adjustment" is the output of the fuzzy system. It is a value ranging from –0.1 to 0.1 and contains five different levels: negatively large, negatively a little, still, positively a little, and positively large. The adjusted gray level value can be expressed by the following equation:

$$G_{new}(x,y) = G_{old}(x,y) * (1+\beta) \quad (1)$$

where $G_{new}(x,y)$ and $G_{old}(x,y)$ represent the new gray level value and the old gray level value of the pixel located on (x,y), respectively. β is the gray level adjustment amount, which is the output of the fuzzy adjustment system.

3. NEURO-FUZZY RECOGNITION APPROACH (NFRA)

The neuro-fuzzy recognition approach (NFRA) conducts an area-based image recognition process. Figure 2 illustrates the flow of NFRA, which contains seven steps (Chen 2001).

Step 1 – Step 2:
Image acquisition is the first step of NFRA. Image data can be acquired using a digital camera and transferred to a computer. The second step is to convert the image to gray scale using image processing software. In order to process in an efficient way, an image is usually converted to gray scale before processing.

Step 3:
After converting the image to gray scale, the illumination value of each pixel can be found. All the pixels in the image are separated into three groups in accordance with their illumination values. Illumination values are between 0 and 1, with 0 the darkest and 1 the brightest. The average illumination values of the three areas will be computed and serve as the input to a pre-trained neural network. Figure 3 illustrates the illumination-based image segmentation.

Step 4:
Once the image segmentation is completed, the three average illumination values of the three areas will be sent to a pre-trained neural network to generate three corresponding threshold values, which range from 0 to 255. The training set for the neural network should be diverse so that the trained neural network would be fault-tolerant. Figure 4 illustrates the neural computing process of the three threshold values.

Step 5:
In this step, the fuzzy adjustment is utilized to adjust the gray level values of the image pixels along the boundaries. The gray level adjustment range is from –10% to +10%. Figure 5 shows the flow of fuzzy adjustment.

Step 6:
In this step, each area is thresholded according to its corresponding threshold value. Pixels with gray level values smaller than the threshold values (i.e., darker) are considered as defects (or rusts in this case), and pixels with gray level values larger than the threshold values (i.e., brighter) are considered as background. Figure 6 depicts the illumination-based thresholding process. In Figure 6, the values in the grayscale image represent the gray level values of pixels. The thresholded image is a binary image, with 0's representing the background and 1's representing the defects (rusts).

Step 7:
When the thresholding of all the three areas is completed, the defects in the image can be recognized and the defect percentage can be calculated by counting the percentage of the defect pixels out of all the pixels in the image. Figure 7 illustrates the defect recognition and calculation.

4. MULTI-RESOLUTION PATTERN CLASSIFICATION (MPC) METHOD

There are two resolution levels involved in the multi-resolution pattern classification (MPC) method: the fine resolution level and the coarse resolution level. The fine resolution level (sometimes called the measurement level) refers to the original image. The coarse resolution level refers to a high-dimensional feature space that is mapped from the fine resolution level (Chang 2000). MPC can be broken down into the following steps:

1. At the fine resolution level, the original image is divided into a number of small image blocks. The number of image blocks affects the resolution of the classified (or clustered) image. The larger the number of image blocks, the better the resolution of the classified image.

2. After the division of an image at the fine resolution level, the features of each image block will be extracted and serve as the basis for image classification. In MPC, the spatial gray level dependence method (SGLDM) and the gray level difference method (GLDM) will be adopted for feature extraction (Chang 2000).

3. When feature extraction for each image block is complete, each image block will be mapped onto a high-dimensional feature space as a vector (which can also be thought of as a point in the high-dimensional feature space), with the number of features equal to the dimension of the vector. All the feature vectors in the high-dimensional feature space form the coarse resolution level. The number of vectors (or points) in the coarse resolution level equals the number of image blocks in the fine resolution level.

4. A clustering algorithm will be applied to the coarse resolution level to classify the feature vectors in the high-dimensional feature space. The nearest mean reclassification (NMR) algorithm will be used for feature vectors clustering in MPC (Fukunaga 1990).

5. When the classification of feature vectors is finished, the classification result will be mapped back to the fine resolution level to further cluster the original image. Feature vectors of the same group in the coarse resolution level will make their corresponding image blocks in the fine resolution level fall in the same group.

Figure 8 illustrates the flow of MPC. The structure of the MPC can be represented by the "multi-resolution pyramid" as shown in Figure 9, where n features were extracted from each image block (Chen 2001; Chang 2000).

5. COMPARISON OF NFRA AND MPC

In this section, NFRA, an artificial intelligence method, was compared with MPC, a statistical method proven to be effective in recognizing rust images (Chang 2000). In Figures 10 and 11, it can be seen that both methods performed effectively on rust images. In Figure 10, NFRA seemed to have better performance, because some rust spots in the upper left corner were not recognized by MPC. In Figure 11, both methods had similar recognition results. From the viewpoint of processing time, MPC, which took 4.5 minutes to process an image, is faster than NFRA, which took 5 minutes to process one. Through comparison to MPC, the performance and effectiveness of NFRA can be proven and justified.

6. CONCLUSIONS

In the construction industry, digital image processing has been experimented for use in defect recognition of steel bridge painting. However, there are still some problems associated with this newly proposed application. Poor quality images remarkably affect the accuracy of this application, and non-uniform illumination is usually the cause.

In order to resolve the problem, an illumination-based image recognition technique combining a neural network and a fuzzy adjustment is proposed. The neuro-fuzzy recognition approach (NFRA) segments an image into three parts in accordance with the illumination of the pixels in the image and thresholds an image based on the three areas. The intelligent learning ability of the neural network is utilized for automatic generation of three threshold values through input of the three average illumination values from the three segmented areas. The fuzzy adjustment is used for smoothing the gray level values of the image pixels along the boundaries.

From the application of NFRA to steel bridge rust images, it can be seen that NFRA works effectively on rust image recognition. Through comparing NFRA to the multi-resolution pattern classification (MPC) method, it was proven that NFRA is an effective approach for rust image recognition. The advantage of NFRA is its fault-tolerance characteristic. The disadvantage of NFRA is the system has to be trained before being used.

In brief, the proposed NFRA provides a new approach utilizing artificial intelligence for image recognition that may lead to automation of steel bridge coating assessment in the near future.

7. REFERENCES

AbdelRazig, Y., Chang, L. M., and Skibniewski, M., Automated Quality Assessment of Constructed Surfaces Through 'Intelligent' Image Processing. *Proceedings of the 16th International Symposium on Automation and Robotics in Construction (ISARC)*, pp 22-24 September 1999, Madrid, Spain (1999).

Abraham, D., Iseley, T., Prasanth, R., and Wirahadikusumah, R., Integrating Sensing Technologies for Underground Utilities Assessment. *Infrastructure Condition Assessment: Art, Science, and Practice. Proceedings of the conference sponsored by the Facilities Management Committee of the Urban Transportation Division of the American Society of Civil Engineers*, 25-27 August 1997, Boston, MA, USA, pp 316-325 (1997).

Chang, Y. C., *Statistical Model for MRF Image Restoration and Segmentation*. Ph.D. thesis, Purdue University, West Lafayette, IN, USA (2000).

Chen, P. H., *Digital Image Recognition Methods for Infra-structure Surface Coating Assessment.* Ph.D. thesis, Purdue University, West Lafayette, IN, USA (2001).

Chen, P. H., and Chang, L. M., Application of Neuro-Fuzzy Thresholding to Bridge Painting Defect Recognition. *Proceedings of the 17th International Symposium on Automation and Robotics in Construction (ISARC)*, 18-20 September 2000, Taipei, Taiwan, pp 1105-1110 (2000).

Croall, I. F., and Mason, J. P., *Industrial Applications of Neural Networks.* Springer-Verlag New York Inc., New York, NY, USA (1992).

Fukunaga, K., *Introduction to Statistical Pattern Recognition.* 2nd Edition, Academic Press (1990).

Table 1. If-Then Rules for Fuzzy Adjustment

IF	Positive Difference	AND	Negative Difference	THEN	Gray Level Adjustment
IF	*Large*	AND	*Large*	THEN	*Still*
	Large		*Medium*		*Positively Small*
	Large		*Small*		*Positively Large*
	Medium		*Large*		*Negatively Small*
	Medium		*Medium*		*Still*
	Medium		*Small*		*Positively Small*
	Small		*Large*		*Negatively Large*
	Small		*Medium*		*Negatively Small*
	Small		*Small*		*Still*

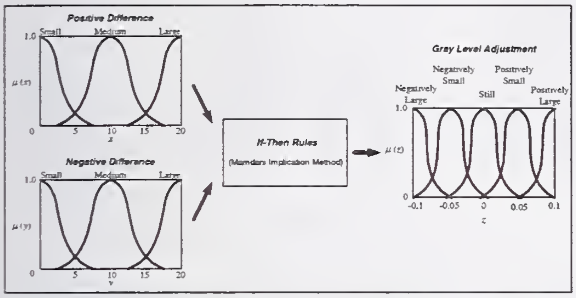


Figure 1. Schematic Representation of Fuzzy Adjustment

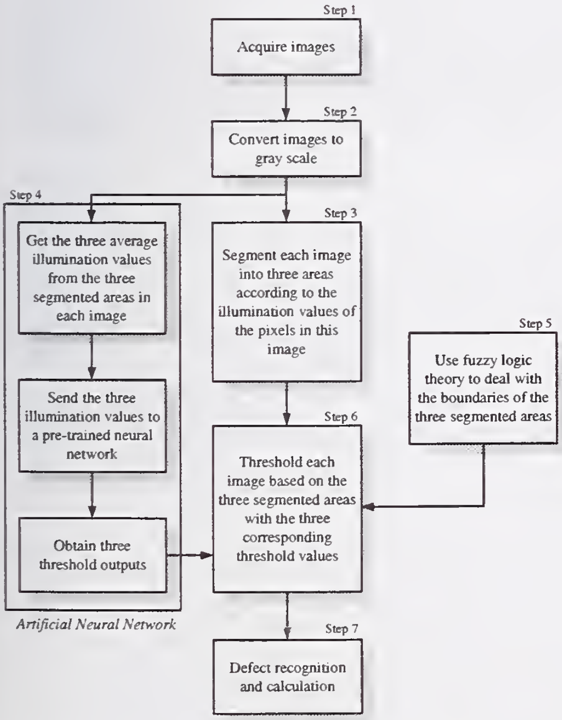


Figure 2. Neuro-Fuzzy Recognition Approach (NFRA)

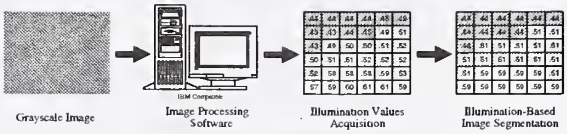


Figure 3. Illumination-Based Image Segmentation

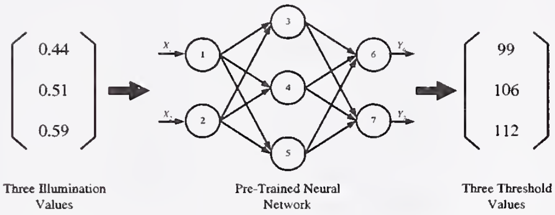


Figure 4. Neural Computing of Threshold Values

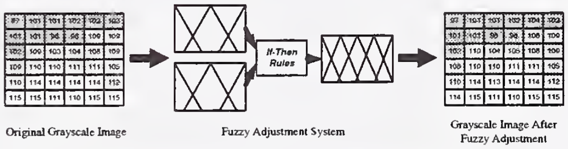


Figure 5. Fuzzy Adjustment on Boundary Pixels

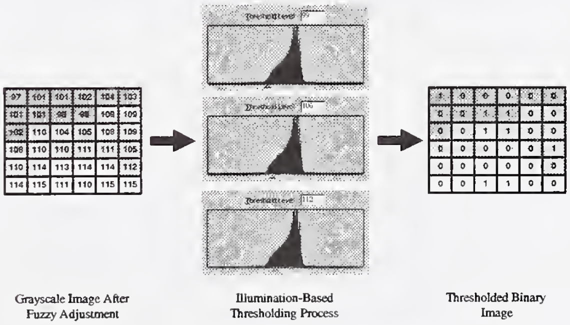


Figure 6. Illumination-Based Thresholding Process

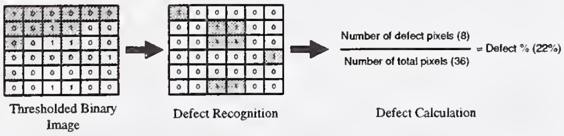


Figure 7. Defect Recognition and Calculation

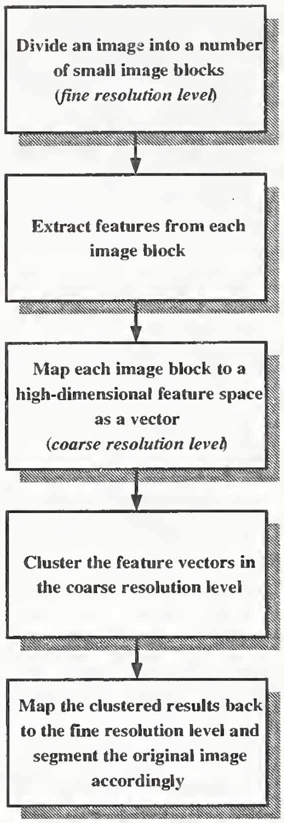


Figure 8. Flow of Multi-Resolution Pattern Classification (MPC) Method

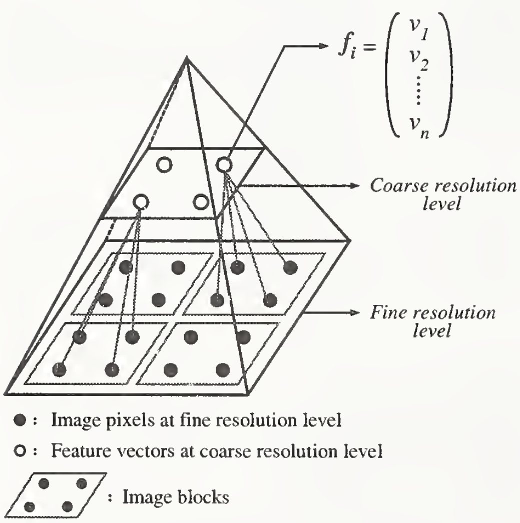


Figure 9. Multi-Resolution Pyramid

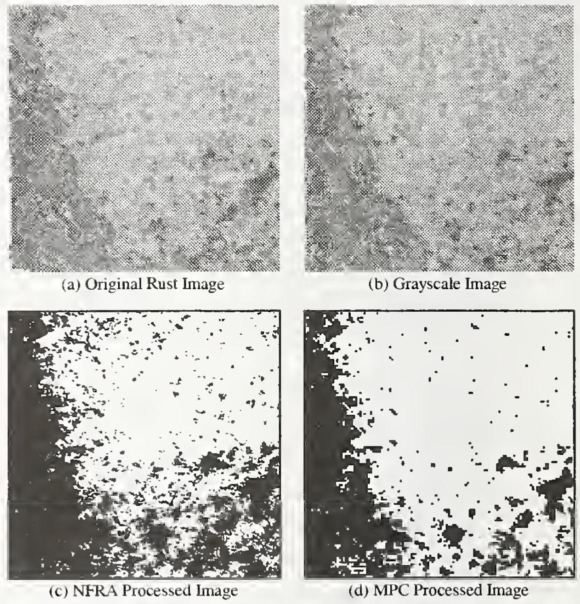


Figure 10. Comparison of NFRA and MPC (I)

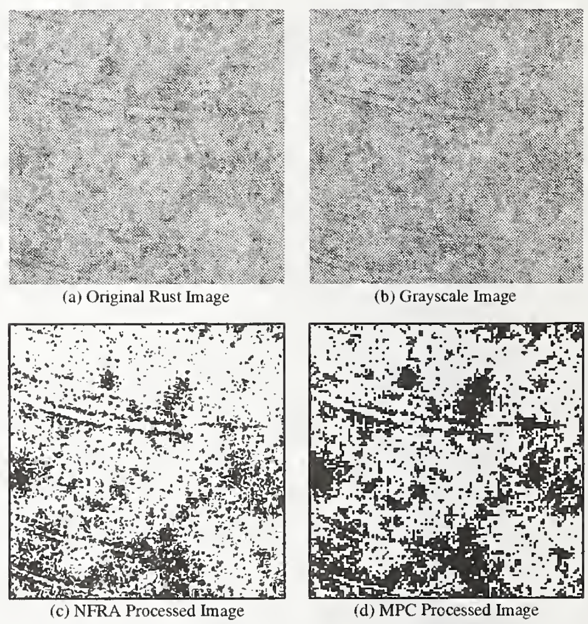


Figure 11. Comparison of NFRA and MPC (II)

THE APPLICATION OF A 3D SCANNER IN THE REPRESENTATION OF BUILDING CONSTRUCTION SITE

Naai-Jung Shih

Professor, Department of Architecture
National Taiwan University of Science and Technology
73, Section 4, Keelung road, Taipei, 106, Taiwan, ROC

Abstract: This study offers a method for creating a representation of an on-going construction site by using a geometric approach. The representation includes the definition of 3D data between the design and construction stages. A method is proposed for using a 3D long-range laser scanner to retrieve information on site occurrences, in order to describe the building construction process based on what really occurs at a site. 3D scans are made in stages or at various time intervals. Retrieved shapes are collected in a discrete manner or in a configuration as a whole. The data collection is conducted wholly or selectively for key referencing based on registration points. Investigated issues involved are target positioning, identification, retrieval, tracking of objects, behavior description, characteristic description, and integration of segmental geometric information. A special procedure is developed to fabricate initialized object parts through a 3D rapid prototyping process to facilitate communication in substantiated form.

Keywords: 3D scanner, construction

Introduction

VR or 3D modeling approaches are used to simulate tasks to facilitate a better visualization of building construction. (Retik et al., 2000; Retik & Shapira, 1999; Vaha et al., 1997). However, making a digital representation of a construction site is a complicated task. Not only are some objects too trivial to be represented, but also the undefined data may come from the construction process itself. Starting from the design stage, 3D computer models are usually used to facilitate an inspection in the design stage only, by showing related component definitions like walls, columns, openings, etc. in a built form. This type of representation may be sufficient to visualize a building in its final stage or be used as the basic definition of process. But an actual construction is much more complicated in terms of objects and associated motions. Not only are the defined activities in the construction schedule more numerous than the original design models can describe, but also the machinery, workers, materials, and all objects presented and not included in the original 3D model are left to be noted only in a text or chart form, based on judgments derived from related experiences.

Construction site monitoring is an on-going process that records and monitors data for immediate and post-construction analysis (Al & Salman, 1985; Atkin, 1986; Bjoerk, 1993). The monitoring of a site and the correspondence of activities defined by a schedule requires object identification and a thorough record of site occurrences. To achieve this goal, the function of 4D monitoring focuses on a pre-construction study for the better management of a site afterward (Haymaker & Fischer, 2001). During construction, object identification and comparison to scheduled activities are essential functions of site monitoring, and are usually conducted by supervisors as human-based tasks. That means site monitoring is an analogical process. Functions that could enable the digital identification of an object and the ability to check whether it is on schedule would be very helpful.

A reversed description method is proposed to improve the disadvantages of 4D technologies by recording site occurrences in geometric form and comparing them with the activities recorded on the schedule. The comparison is made as a reference base for evaluating actual progress. In contrast to 4D technology, which is based on data from the design stage, the application of a 3D scanner retrieves data on a construction site to facilitate an as-built description of models in a precise and reversed verification manner.

Proposes

This study proposes a method for represent an on-going construction site through a geometric approach. The representation includes the definition of 3D data between the design and construction stages. A method is proposed for using a 3D long-range laser scanner to retrieve information on site occurrences, in order to describe the building construction process based on what really occurs at a site. 3D scans are made in stages or at various time intervals. Retrieved shapes are collected in a discrete manner or in a configuration as a whole. The data collection can be conducted wholly or selectively for key referencing based on registration points.

Systems

The system consists of a long-range (50-100 meters) 3D laser scanner Cyrax 2500 (see Fig. 1) that comes with an editing software, Cyclone 3.1. The laser can create a matrix of point clouds up to 999*999 dots in width and height. Scans can be made individually or registered onto a large project by referring to tie-points. Each scan can reach a tolerance of 2mm / 50 meter. The system comes with a notebook computer, which figures 1GHz CPU and 512 MB RAM, to handle the data received on site.

Levels of manipulation

The possibilities for manipulation of scanned geometric data can be classified into three levels:

- Level I – point clouds: Initial data retrieval is collected through 3D scans. Each point is initialized with x-, y-, and z-coordinates only. Points can either be initialized with absolute coordinates or modified to show relative location. At this level of three-dimensional representation, measurements can be made to estimate linear distance between two points or between a point and a plane. Data are in a generic form that opens up a referencing base for following manipulation. It is suggested that at least one copy of raw scan data should be kept unchanged in terms of point coordinates, viewpoint, and the images associated with scan orientations.
- Level II – object initialization: Structural detail, which represents the components and their mutual relationships, is created by fetching or matching points with geometries like mesh or various shapes of primitives. In contrast to one-dimensional representation, volumetric description enables calculations to be made of surface area and of object volume. Relative location between objects becomes more meaningful for the boundary of objects is defined and can be manipulated afterward. In addition, the volumetric data can be exported to domain-specific applications for further visualization or analysis.
- Level III – attribute designation: Initialized geometry objects are mapped with images or textures to illustrate surface attributes. At this level, additional visual detail is presented. Models created at previous levels can be used to build up a more visually attractive scene.

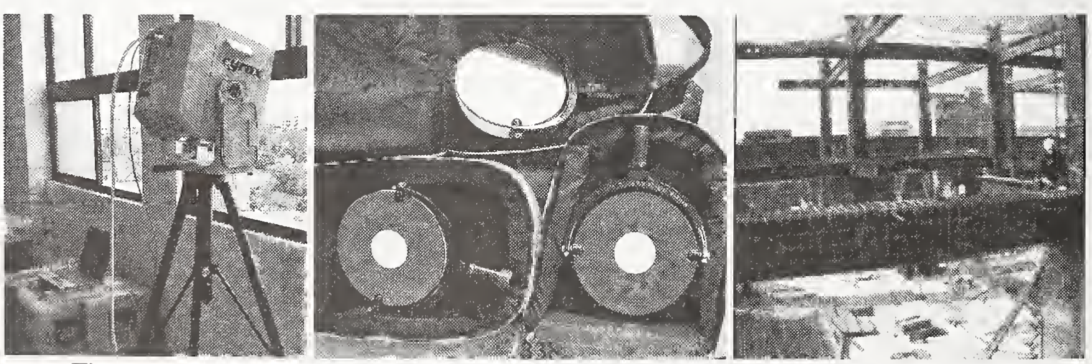


Figure 1. Cyrax 2500 3D laser scan system, registration points, and site example

Object representation

Why do we need a detailed geometric description of an object at a construction site, especially when the data retrieved can be too fragmental to facilitate further manipulation? The reason is a construction site is a very complicated scene; a detailed description of it can facilitates analysis without the risk of missing any valuable data. The fragmental nature of site information can now be changed by registering the fragments together.

One of the 4D disadvantages actually comes from the effectiveness of digital representation of site occurrences. In general, two types of comparison exist between real objects in physical form and digital form:

- An object's digital representation: This representation comes from scanning or computer modeling of real objects. Computer models for design evaluation are usually created before construction begins. Objects initialized from scanned point clouds represent a physical form in a finished stage. Nevertheless, the intermediate stages between design and construction stages can also be represented through scans of partially completed building parts and components allocated accordingly.
- An object's behavior and schedule: The big picture of a digital construction representation includes construction components, component behavior, and component relationship to the construction schedule. The representation of a construction at a certain stage does not necessary reveal the behavior of its components. This behavior, as described in a previous section of this paper, provides more details than the information collected for a certain task.

Process and Methods

To match objects and their roles specified in construction schedule, there are many sub-tasks have to be conducted. Manipulation geometric and image data are classified into three levels of data manipulation:

- dimension-related geometric characteristics check: The initial level of data manipulation emphasizes a preliminary use of scanned data and dimension-related checks based on geometric characteristics. Usage of scanned data can be shown in analysis, measurement, record, or visualization. Measured data can be used for dimensioning, defining relative location, and recording geometric atributes. For example, a scan records geometric information of a steel joint for further control such as spacing or form works. Steel-bar-related checks that can be derived from the scanned data include spacing, size, numbers, and the relationship with adjacent components. Dimension-related checks are made in terms of standard deviation, which mainly comes from stress-related (pressure, force, fire) deflection or bad construction quality. Geometry-related attributes of the scanned data are categorized based on their usage in one, two, and three dimensions.
- object identification and behavior monitoring: During construction, when objects are scanned and initialized into geometries, there is a need to tell how close the scanned geometries are to their product specification. The identification can be categorized into image-based and geometry-based approaches. Images usually contain more details than geometric representation. While the function of abstraction and segmentation for images or videos is limited, a pre-design study in a visual form must be conducted from a modeling approach, leaving an image approach to be conducted based upon what is already constructed as an as-it-is situation.

Behavior represents the spatial movement of construction components over a period of time. An object's movement is the response of its own or external forces along a time span. The monitoring of behavior notifies us of an object's activities. With the environment scan included, the object's response to its surroundings is also specified from itself as well as construction instruments.

- integration of segmental geometric information: Due to unexpected blockage of other objects, scanworlds have to be integrated to complete an omni geometric representation of an object. Scanned data or image-derived photogrammetric data offer partially complete descriptions. Only when an object's movement or orientation changes so as to lead to more exposure to the receiving devices or applications can additional data be perceived as a complete set. In a chronicle recording, each object may inherit a different level of data segmentation. Integrating the chronicle-based segmentation is a challenging task and is currently made by registration through referencing points.

Exemplification of 3D scans has been made to construction site for dimension-related check of geometric attributes. Related studies are also made to interior design (proofing; as-built scan for renovation documenting) and a long term chronicle documenting of construction in the mean time.

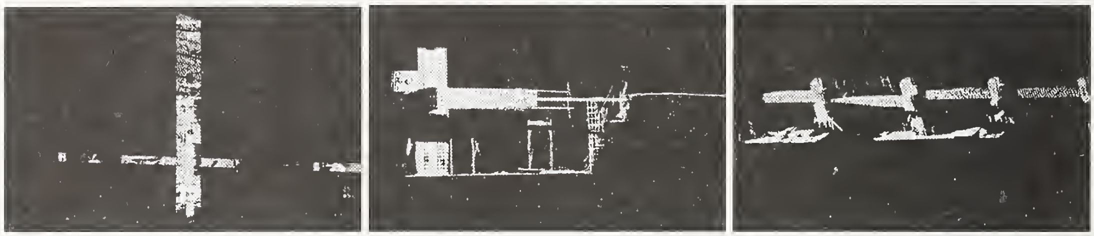


Figure 2. Two referencing sections through whole site

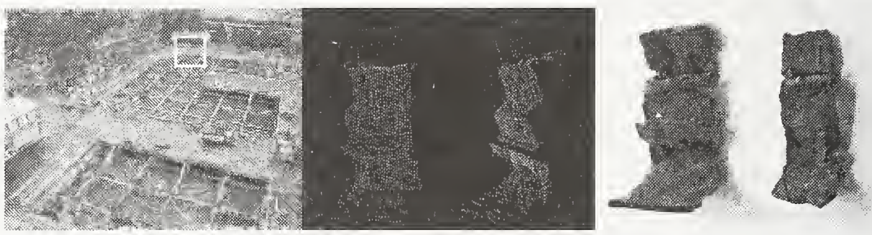


Figure 3. Photo, point cloud, and RP output (from left to right) of retaining wall section

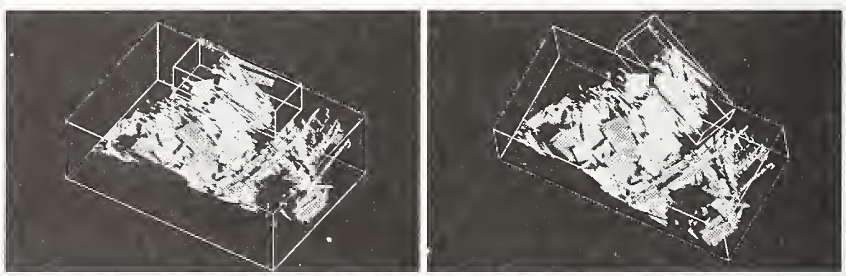


Figure 4. Two segmented clouds are shown in ordinary (left) and minimum volume (right) presentation

Figure 5. Center parts of the cloud is segmented into two and stored in difference layers

Section referencing through whole site

A 3D scan is suitable for a clear review of a cross-section of a site (see Fig. 2), not just to review an object as a whole from outside. The process does not have to be undertaken in a traditional manner, made by creating models and conducting software cuts. The purposes of the scan are to show:

- structural and compositional relationship: to reveal aggregation and segregation of building parts
- hidden components and spaces: to illustrate what can be left unseen
- part-whole generating process: to create a time-based transformation description of the "how & why" of occurrences

Most important of all, sections are made based on what really occurs at a site. The section is a good manner of spatial representation suitable for construction quality control, such as for levels or clearance between slabs. Additionally, initialized object parts can be fabricated through a 3D rapid prototyping process to facilitate communication with substantiated forms (see Fig. 3).

Boundary and box approximation

The number and complexity of objects at a construction site may be too numerous and detailed to be modeled exactly. But recording the presence of objects is still an important task to be aware of. In order to conduct the following study, a boundary or box approximation is conducted for the following purposes:

- simplified volumetric representation: The boundary representation can be created based on either current orientation or minimum volume illustrated by a rectangular wire frame (see Fig. 4).
- interference analysis: The presence of a boundary enables an interference check between adjacent objects in a user defined manner. For example, a designer or site supervisor can select the segmented clouds of interest to create a group, and then compare the boundary with another group for interference. This method enhances current interference checks provided by similar 3D modeling applications in a systematic review of, for example, a plumbing part and an HVAC part. Not only does the analysis provide dimension-related data, but also a diagrammatic illustration of the field of influence.
- data segmentation: In each scan, all points in a cloud are considered as an entity (see Fig. 5). Although intensity-based segmentation can be conducted, the result subjects to the distribution of points and the density may not provide a segmentation that matches the perception of ordinary experience. Cloud segmentation is usually manually made and interpreted as needed.

Trade-off of 3D scan

- Time-saving operation: The data retrieval is directly made by a scan that takes about 17 minutes for each scanworld. Even with the follow-up object initialization included, the time needed is much less than for conducting computer modeling from scratch.
- Effort-saving operation: The geometric data are retrieved with x-, y-, and z-coordinates. The effort required to process geometric data is much less than for modeling that starts from each component.
- Spatial barrier free: Scans can be conducted from different orientations to retrieve almost all the part of objects, without being presented close to the parts. A scan of a distance also helps to prevent possible dangers in positioning oneself above ground level.
- The trade-offs of the 3D scan approach include:
- Tolerance: Scanned data still bears tolerance in terms of position, distance, or angles per unit length.
- Data amount: When a scan is made of 999x999 points, the registration of several scanworlds can easily accumulate up to millions of points. The object initialization and differentiation of scanned data can still be effort consuming. One way to simplify the task is to select key objects of interest only or to apply box approximation representation.

Conclusion

This study offers a method for representing a site under construction through the data retrieved from a 3D scanner. The data are point clouds that can be stored as records or can be initialized as geometry shapes afterward for the purpose of analysis and visualization. Although trade-offs exist, a site under construction can now be digitized. In contrast to the traditional 4D approach, the digitized data provide a referencing base made after the design stage for forward verification and backward simulation.

REFERENCES

Al, S. and Salman, T.: 1985, *Large-scale construction projects: management, design and execution*, London: Batsford.

Atkin, B.: 1986, Computer supported information interchange for construction projects, p.55-65, *IABSE workshop* Zuerich.

Bjoerk, B.C.: Autumn 1993, A case study of a national building industry strategy for computer integrated construction, *International Journal of Construction Information Technology.*

Haymaker, J. and Fischer, M., "Challenges and Benefits of 4D Modeling on the Walt Disney Concert Hall Project," CIFE working paper #064, January 18, 2001.

Retik, A., Clark, N., Fryer, R.: 2000, Hasrdiman, R., McGregor, D., Mair, G., Retik, N., and Revie K., *Mobile Hybrid Virtual Reality and Telepresence for Planning and Monitoring of Engineering Projects*, http://www.brunel.ac.uk/depts/mes/Research/Groups/vvr/vrsig97/proceed/018/VRSIG-TX.HTM.

Retik, A. and Shapira, A.: 1999, VR-based planning of construction site activities, 8(6), pp. 671-680, Special issue: *Organisation and management of construction.*

Shih, N.J.; Tsai, Y.L.; Lai, Daniel, and Chang, H.Y., " A Preliminary Application of Panoramic Supervision System in Construction Site," Proceedings of the 2001 Information Visualization, pp. 103-8, July 25-27, 2001, London, England (2001).

Vaha, P., Kansala, K., Heikkila, R. and Kaisto, I.: 1997, Use of 3-D product models in construction process automation, *Automation in construction*, 6(2), pp. 69-76.

DATA INTERPRETATION FROM LEUZE ROTOSCAN SENSOR FOR ROBOT LOCALISATION AND ENVIRONMENT MAPPING.

by

D. W. Seward[1], S. D. Quayle[2], K. Zied[3], C. Pace[4].

ABSTRACT: Two applications for the use of a laser-scanning device are currently under investigation at Lancaster University. Lancaster University Computerised Intelligent Excavator (LUCIE) is an autonomous excavator which navigates using GPS and compass readings. Work is currently concentrating on navigational safety, for which the rotoscan sensor is employed for obstacle detection, and for possible self-localisation and environment comprehension in ambiguous operational states. Starlifter is a robotic arm built by Construction Robotics Ltd. The rotoscan sensor in this instance is to be mounted on the tool head and used as a final positioning navigation tool. Both these applications rely heavily on the interpretation of the received data, and the ability to filter out any interference. This paper initially outlines the mode of utilisation of the laser range finder within such applications and then proceeds to investigate the implications and potential limitations of such a sensor following the analysis of the sensory data from external field trials.

KEYWORDS: collision detection; navigation; robotics; sensors; surface estimation.

1.0 INTRODUCTION

Construction robots have been under development for many years, although in the field they have not yet fulfilled their potential [1]. Although it is relatively straightforward to get the robot to achieve its coarse objectives, it is considerably more difficult to ensure that the task is completed safely and accurately. One of the ways for handling safety is to fence the robot off away from human beings, and other disturbances, however this will often remove the gains of using a robot in the first place. The alternative approach is to have the robot respond predictably to its changing external environment. To do this effectively the robot needs to accurately sense its surroundings. Accurate sensing is also critical for tasks requiring accurate positioning and operations.

With regards to safety, there are numerous safety concerns that need to be catered for in the safety validation of such a system. Apart from internal operational integrity, of fundamental importance in such systems is the need to ensure correct interaction between the autonomous system and the environment. The correctness of such interaction will be dependent on the autonomous system's perception of its surroundings, these being dependent in turn on the exteroceptive sensory suite of the system, of which the laser scanner is a critical element.

Correct interpretation of the range sensing data is also critical for accurate positioning operations. In this case, accuracy becomes a critical parameter.

However, the operation of the laser scanner as a range sensing device gives rise to substantial

[1] Engineering Department, Lancaster University, Lancaster, LA1 4YR, U.K., Tel: +44 1524 593010, Email: d.seward@lancaster.ac.uk
[2] Engineering Department, Lancaster University, Lancaster, LA1 4YR, U.K., Tel: +44 1524 593095, Email: s.quayle@lancaster.ac.uk
[3] Engineering Department, Lancaster University, Lancaster, LA1 4YR, U.K., Tel: +44 1524 593824, Email k.zied@lancaster.ac.uk
[4] Department of Manufacturing Engineering, University of Malta, MSIDA MSD06, Malta, Tel: +356 32902058, Email cmpace@eng.um.edu.mt

ambiguity in the interpretation of the data, partly because of the limited information provided by scanning only in one plane at a time and mainly due to inherent limitations of the sensory system, such as dependency on object reflectivity and other physical parameters, which gives rise to uncertainty in the interpretation of the sensory information.

Knowledge of such limitations in the information quality from such a sensor is a fundamental necessity in developing a system that is capable of managing and utilising the information adequately for the task at hand. The focus therefore in such an application is to better comprehend;

1. the application of the rotoscan itself as an additional verification tool for other exteroceptive sensory data.
2. The inherent limitations in the interpretation of the rotoscan data for obstacle recognition, localisation, and accurate measurement.

This paper thus looks at aspects of concern in the use of such a sensing device for the underlying applications, given test runs that were performed for the sensor.

2.0 APPLICATION AREAS

The Rotoscan RS3 range sensor under test is a twin laser beam unit, with each beam scanning 90° leading to a 180° flat plane of view. The range of detection is up to 25m with obstacles as small as 70mm being detectable. Two protection boundaries can be programmed into the sensor, an object boundary and personal boundary. Both boundaries trigger relays once broken. Alternatively a serial data string is available.

Two application areas are being considered for the sensor. The first is Starlifter which aims to combine the laser scanner data with a camera view to enable accurate final positioning, the second is Lancaster University's Computerised Intelligent Excavator (LUCIE) which uses the sensor for navigation and operational safety.

2.1 Starlifter robot

Starlifter is a hydraulically powered portable robot with six revolute joints. The joints can be simultaneously locked in any position with power and control shut down to provide a stable platform to deploy heavy tools up to 200 kg at any orientation. Starlifter is currently configured to carry diamond core drilling tools and concrete saws for construction applications - for further details see reference [2].

One of the uses of Starlifter is working in hazardous environments in which human involvement is eliminated. In this case the robot is located in a remote position from the operator. Consequently information about the position of the robot base relative to the working area is required to enable the operator to position the robot accurately and safely.

An important use of the range scanner with Starlifter is to assist in the positioning of the robot base relative to the working area so that the robot can operate correctly within its working envelope. In this case the laser scanner is attached to the robot base, which is kept level by using balancing rams. The scanner's object safe field can then be adjusted for collision avoidance. Working area perception is improved via the use of a vision system.

2.2 LUCIE

LUCIE is a JCB801 retrofitted with Danfoss electro-hydraulic valves and three individual PC104 units to act as Low Level Controller, Activities Manager and Safety Manager [1]. The range sensor is mounted approximately 0.5m high on the excavator cab. This allows a full field of vision around the side of the excavator for slewing operations, as well as the area immediately around the arm and bucket. For final implementation a second range sensor would be placed on the opposite corner of the vehicle giving a swept coverage of approximately 300°. The area directly behind the excavator is of lesser importance and can be protected by other sensing means, whilst the application of the frontal overlap removes blind spots due to operation of the arm.

The investigation in the case of LUCIE focuses on the evaluation of safety aspects and the underlying causes of operational risks for an autonomous excavator in a construction site environment [3].

2.3 Sensor Data Processing

The serial data from the laser range sensor is extracted and processed using a user interface specially developed using the LabVIEW graphical programming environment. Figure 1 shows the front panel of the user interface. The serial data includes start and end identifiers, the scanner status bits and the user data. The user data represent the measured distances in mm per 2 degrees of the 180-degree range. A successive matrix manipulation of the serial data is employed to separate the user data and the status bits in the data processing section of the code diagram. Further processing of the user data is performed to visualise the measured distances in a Cartesian or polar co-ordinate graph.

The processed data can then be stored for further off-line evaluation. The data obtained represents the co-ordinates of a horizontal plane passing through the laser scanner known as a segment. Moving the laser scanner up and down at different levels, or by tilting the scanner, and by feeding back the vertical or tilt position of the scanner a three-dimensional graph of the working area can be constructed. Figure 1 illustrates a Cartesian map creation by stacking a number of scans taken by shifting the scanner in a vertical plane.

The data is also stored for post processing purposes for further analytical work. The processing and post-processing of the data and the visual representation provide the necessary tools to determine the operational characteristics as seen in the experimental tests and conclusions that follow.

3.0 EXPERIMENTATION

A series of tests were carried out with the range sensor to determine the performance characteristics in external environments and under the likely operational conditions to which the sensor will be subjected.

The objective of the tests was to determine the reading reliability in external environments, particularly:

i. when the scanner is utilised for the detection of surfaces that are irregular and with poor reflective properties
ii. When the scanner is attached to a moving platform and driven over rough terrain, and is therefore subject to machine vibrations and sudden displacement changes.

From the results obtained the level of accuracy of the data, given the specific operating conditions, was to be determined. In this manner, the interpretation of the data during operation could be adapted according to the sensor characteristics.

Experimental runs consisted of both static and dynamic tests, with 180° scans being recorded every 0.75sec, The total number of scans per test ranged from 50 to 80. The distance measurements were extracted through the sensor data processing program and then post-processed to extract statistical characteristics of the data.

Final tests were performed on the plotted data for surface extraction via the Iterative End Point Fit algorithm [4], to determine the portability of such a routine, given the quality of the readings obtained in external environments.

Figure 2 illustrates a typical test run for the laser scanner, where the scanner is placed on a mobile platform and with the motion of the scanner as indicated.

3.1 Static Testing

Static tests were carried out with the laser scanner stationary and pointing to a static or partly static environment (i.e. with random object presence). Figure 3 illustrates the scanned data for a completely static test with figure 4 illustrating the mean and standard deviation values for the test. The number of scans recorded for this test was 50.

From the graph it is immediately appreciated that as distances from the scanner increase the standard deviation for the readings taken over the total number of scans increases. Given the prior measurements from the scanner of the detected surfaces, the readings give a zero mean error for the vast majority of the readings. Non-zero mean errors tend to occur mostly on poor reflective surfaces, since the surface is not adequately detectable.

The rate of change of standard deviation though, is relatively irregular and depends greatly on the type of surface being detected. This can be immediately noted from the detection of vehicles that cause unexpected fluctuations in readings due to the type of surfaces (including glass) on which the laser beam impinges. Yet still, an exponential relationship between distance and reading variance has been found to suitably fit the vast majority of readings, as would be expected from the general class of range sensor models described by Elfes [5].

Spurious readings have also been noted to cause unexpected increases in standard deviation readings, particularly at edges of surfaces and mostly at relatively distant surfaces. This is likely to be caused due to the possible repeatability errors in the range sensor output.

In the case of tests involving partial dynamic features in the environment, where an object was placed within the identified zone for a relatively short period of time at random, it was noted that the temporary presence of the object mainly causes a substantial increase in standard deviation in the sensory readings with a minor effect on mean value. The mean is found to only drop slightly in value as expected due to the presence of shorter distance measurements within the time interval where obstacles are present. Similar results were obtained when simulating rain conditions with standard deviation readings increasing whilst still maintaining relatively constant mean values.

3.2 Dynamic Testing

Dynamic tests were principally carried out with the laser scanner attached to a mobile platform and moved over relatively rough terrain, inducing minor vibrations to the sensor. Figure 5 outlines the results from such a dynamic test for the example illustrated in figure 2. In this specific test, the laser scanner is brought close to the objects at a relatively linear and constant speed.

To determine the repeatability of the sensory data given such operating conditions, successive readings were mapped onto each other following an angular and linear translation, so as to give the least mean square error between successive readings. The least mean square error is given in figure 5(i) as the test proceeds for each scan taken (a total of 55 scans were taken in this test). Figure 5(ii) and (iii) outline the estimated linear and angular translations to obtain the least mean square error.

From the graphs it is easily noticeable that the root mean square error is much greater than the standard deviation readings obtained for the static readings. Most of this increase in the readings' variance can be said to be due to the motion of the sensor and the induced vibrations. Indeed towards the end of the test with the sensor's velocity almost zero, root mean square error values drop down to values close to those obtained for standard deviation in the static tests. This clearly indicates a substantial limitation of the laser scanner for accurate distance readings during motion, where sensor variance increases as expected due to the errors induced from the motion vibrations.

It can also be noted that as distance measurements increase a higher mean square error is observed with the range sensor in motion, as would be expected given the results of the static tests. This can be noticed from figure 5(i) where there is a gradual drop in root mean square error values as the laser scanner is brought closer to the obstacles. This drop is roughly exponential in nature, as would be expected.

The root mean square error has also been noted to increase with speed for the same distance measurements. Again, this would be expected partly because of the increased 'disturbance' between two successive scans.

3.3 Surface Estimation

The Iterative End Point Fit algorithm was applied to each and every set of test scans to determine the performance of the algorithm given the characteristics of the tests. The algorithm was found to perform poorly when applying the plotted readings directly, particularly under conditions where the variance between successive readings was large. This resulted in totally different surface profiles being generated between successive scans. Better results were obtained for tests with reduced variance in the readings, although spurious readings did cause sudden changes in the estimated surface profile.

Improvements to the algorithm were obtained by introducing a filtering algorithm to eliminate spurious readings between successive scans and by averaging multiple readings to smooth out any ambiguities within single scans.

However, due to the nature of the surfaces being scanned, the algorithm still gave erroneous interpretations when not scanning large flat surfaces, and therefore was not found to be adequate in its simple form for external environment surface recognition.

4.0 CONCLUSIONS:

The tests outline a number of interesting aspects with regards to the use of the Rotoscan range sensor in the outlined applications.

Primarily it is seen that the reliability and accuracy of the rotoscan readings degrades with the motion of the range sensor, mostly as a result of the induced irregular motion from the terrain characteristics. However, this does not diminish to any major extent the applicability of the sensor given that the measurements are made beyond the immediate vicinity of the autonomous system. Measurements made close to the autonomous system, requiring high accuracy, , cannot be made while the scanner is in motion. Accurate close-up measurements therefore require a stationary platform, which results in a drastic reduction in variance. This reduction in variance has also been found to be related exponentially with distance, and this exponential behaviour has been found to occur both during stationary and dynamic scans.

The range sensor has been noted to be ideal for detecting uniform surfaces perpendicular to the scanning plane. Slopes and other obstacles with uneven surfaces though detectable are much more difficult to identify, and distinguish from sensor errors. Indeed, the detection of such surfaces also induces larger variance values than for flat surfaces. In addition the ability to distinguish features as distance increase drops drastically, particularly if such features are of an irregular nature (i.e. not a flat perpendicular surface). It was also noted that transparent materials such as glass and water (rain) are not detectable to any significant extent, resulting mostly in spurious readings. However spurious readings seem to occur even in the absence of such surfaces and at a rate greater than for internal environments.

With regards to the iterative end-point fit algorithm for feature extraction, as stated earlier on, the algorithm was found to act poorly and was only useful in identifying large flat, perpendicular surfaces. The nature of the readings and the irregularity of the surfaces tend to inhibit the correct identification of the readings given the relatively high variance that occurs in external environments.

The above outline the main limitations for the use of the laser range finder in external environments. The sensor has been found to perform suitably particularly if relatively rough estimates of distance measurements are required, particularly when the scanner is in motion. The relatively high variance in the readings may not be too much of a hindrance if the data is only required to roughly estimate the distance from obstacles. In addition, the variance tends to drop as obstacles get closer to the autonomous system.

However, the nature of the readings, inhibits proper identification of environmental features unless, the sensor is stationary and multiple readings can be taken. In addition, feature identification and consequently self-localisation requires the presence of regular surfaces that are distinguishable for the relative clutter in the data caused by other irregular and poorly reflective surfaces. The absence of such type of surfaces and adequate operating conditions is highly likely to inhibit the correct application of the sensor for more accurate identification tasks.

5.0 REFERENCES:

[1] D. W. Seward, S. D. Quayle 'System Architectures and Safety for Mobile Construction Robot', Proceedings of the 13th ISARC, Tokyo, Japan, 1996.

[2] K. Zied, D. Seward, J.Riehl, A. Dolman: 'The Development of a Robotic System for Tool Deployment in Hazardous Environments', Proceedings of the 17th ISARC, Taiwan, 2000.

[3] C. Pace, D. W. Seward ' A Safety Integrated Architecture for an Autonomous Excavator', Proceedings of the 17th ISARC, Taiwan, 2000.
[4] K. Diath, A. Siadat, M. Dufaut, D. Wolf, 'Navigation of a mobile robot by locally optimal trajectories', Robotica, Vol17, pp553-562, 1999.
[5] A. Elfes 'Multi-source Spatial Data Fusion using Bayesian Reasoning', in 'Data Fusion in Robotics and Machine Intelligence', Mongi A. Abidi, R. C. Gonzales (Ed.), Academic Press, 1992.

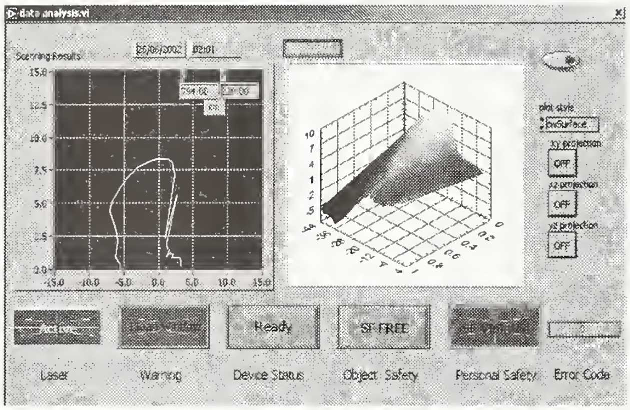


Figure 1. The front panel of the graphical user interface of the laser scanner

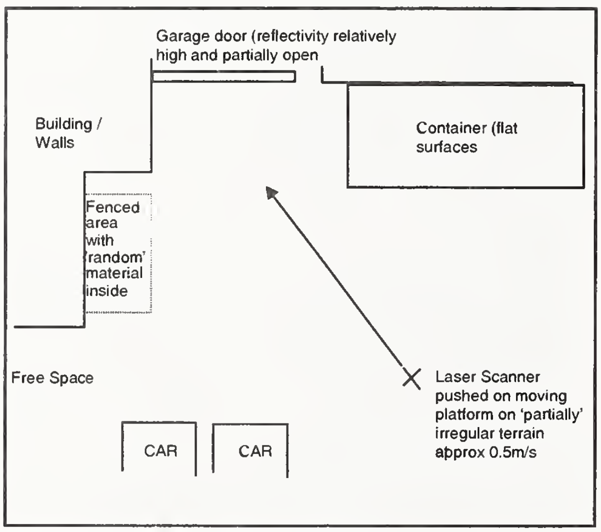


Figure 2. Typical test layout for Laser Range Sensor

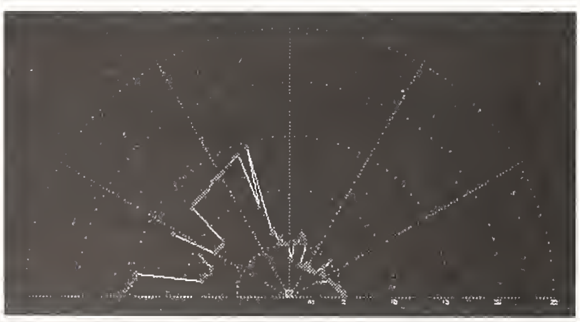


Figure 3. Static sample scan test in polar coordinates

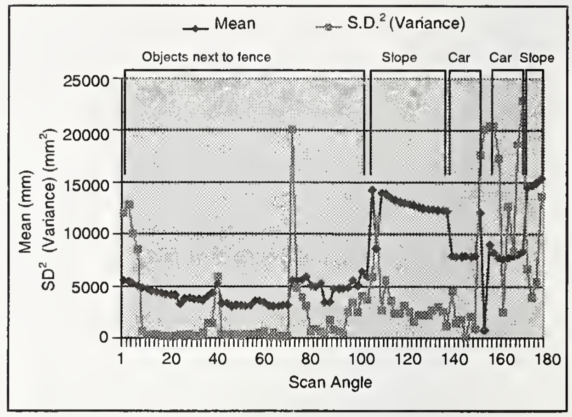


Figure 4. Statistical data (mean and standard deviation) for static test of figure 3.

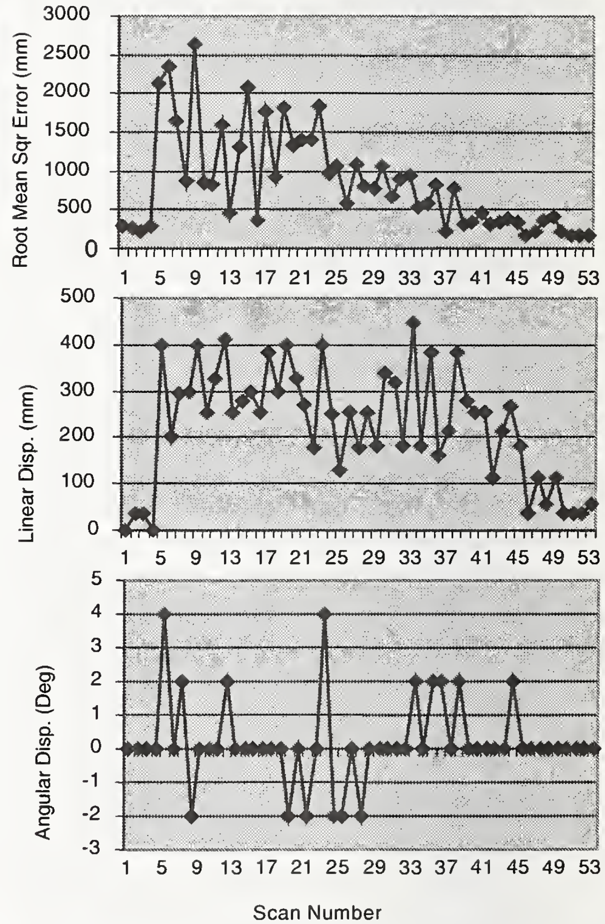


Figure 5. Results for dynamic test in figure D. (i) Root mean square error for every scan sample, (ii) estimated incremental linear displacement, (iii) estimated incremental angular displacement

Experiences with Point Cloud Registration [1]

by

Christoph Witzgall[2] and Geraldine S. Cheok[3]

ABSTRACT: The development of LADAR (laser distance and ranging) technology to acquire 3D spatial data made it possible to create 3D models of complex objects. Because an unobstructed line-of-sight is required to capture a point on an object, an individual LADAR scan may acquire only a partial 3D image, and several scans from different vantage points are needed for complete coverage of the object. As a result there is a need for software which registers various scans to a common coordinate frame. NIST is investigating direct optimization as an approach to numerically registering 3D LADAR data without utilizing fiduciary points or matching features. The primary capability is to register a point cloud to a triangulated surface - a "TIN" surface. If a point cloud is to be registered against another point cloud, then the first point cloud is meshed in order to create a triangulated surface against which to register the second point cloud. The direct optimization approach to registration depends on the choice of the measure-of-fit to quantify the extent to which the point cloud differs from the surface in areas of overlap. Two such measures-of-fit have been implemented. Data for an experimental evaluation were collected by scanning a box, and registration accuracy was gauged based on comparisons of the volume and height to known values.

KEYWORDS: LADAR; measures-of-fit; point cloud; registration; TIN; triangular mesh.

1. INTRODUCTION

In recent years, the National Institute of Standards and Technology (NIST) has investigated metrological aspects of LADAR (Laser Distance and Ranging) scanning, addressing both hardware and software issues [2,4,5]. Hardware calibration issues include statistics for direct range measurements, their dependence on target color, distance, and angle of incidence. Corresponding experiments will be the topic of a forthcoming report [2]. In this work, the focus is on software issues, in particular, the registration of LADAR scans taken of the same scene from different vantage points. Statistical experiments aimed at assessing triangular meshing for surface modeling are described in another forthcoming report [5].

The data for the software experiments consist of four LADAR scans taken indoors of a

$$0.914\,m \times 1.219\,m \times 1.524\,m$$
$$(3\,ft \times 4\,ft \times 5\,ft)$$

wooden box. The task is to create a triangular mesh representing the box and surrounding floor. The accuracy of that representation is gauged by calculating from it the volume and height of the box. The accuracy of the box dimensions is ± 1.58 cm (1/16 in) and assuming worst case, the volume error is ± 0.4 % [4].

The registration methods considered here match a point cloud against a triangulated elevated surface. The methods are conceived as optimization problems, which apply rigid trans-

---

[1] Official contribution of the National Institute of Standards and Technology (NIST); not subject to copyright in the United States.

[2] NIST, ITL, Mathematics and Computational Sciences Division, Mail Stop 8910, Gaithersburg, MD 20899-8910, witzgall@nist.gov

[3] NIST, BFRL, Construction Metrology and Automation Group, Mail Stop 8611, Gaithersburg, MD 20899-8611, cheok@nist.gov

formations to a point cloud in order to minimize a chosen "measure-of-fit", which quantifies the divergence of the point cloud from the surface. Several distinct measures-of-fit will be identified, and the performance of two of them will be examined using the box data. Evaluations of additional options are in progress.

Direct minimization of a selected measure-of-fit differs from the more commonly used ICP (Iterative Closest Point) approach [1,3]. Direct minimization, however, is suited to the purpose of examining the effects of the choice of measures-of-fit.

In a first series of experiments, each of the four scans is registered individually to an exact triangulated model of the box and surrounding floor space. The registered four point clouds are combined into a single point cloud, which is then "cleaned" and meshed to create a surface model of the box. In a second series of experiments, the recreation of the box is attempted by pair-wise registration of the point clouds against themselves without benefit of an accurate reference. The four point clouds are examined in clockwise rotation. A triangulated surface is determined by cleaning and meshing the first point cloud. The second cloud is then registered against the surface model of the first cloud. Once that common frame has been established, the two point clouds are combined, cleaned and meshed to generate a common surface representation of the first two point clouds, against which to register the third point cloud. The process is then repeated by registering the fourth point cloud against the surface representation of the other three point clouds. Finally, all four point clouds -- now registered to the frame of the first -- are combined, cleaned, and meshed to produce a model of the box.

Earlier experiments have been conducted along the same lines [4]. Here we report the outcomes of fully automated procedures.

2. MEASURES-OF-FIT

There are two distinct aspects to distance based measures-of-fit to be identified in this report:

- The definition of generic point-to-surface distance
- The selection of a "norm" by which to distill a single number from many instances of such distances

Commonly considered norms are the

MAX, RMS, and ASD norms.

The MAX norm assigns the maximum absolute value or "size". The RMS ("root-mean-square") norm averages the sum of squares before taking the square root. The ASD ("average-size-deviation") represents the average of the absolute values. Norms typically provide error estimates in the form of deviations from zero. The MAX norm will not be considered here, since it is too dependent on "outliers". Each of the other two norms may be combined with a generic point-to-surface distance measure to arrive at a measure-of-fit.

Three generic distance measures are considered:

- Vertical distance
- Euclidean distance
- Ray-directed distance

A surface is "elevated" or "2.5 D" if it has a unique projection into the x,y-footprint plane. The term "TINsurface" (TIN=triangulated irregular network) is frequently used for triangulated surfaces that are also elevated, and only such surfaces are considered in this report.

"Vertical distance" is defined as the absolute value of the "residual" $r_i = z_i - \hat{z}_i$ of a point $p_i = (x_i, y_i, z_i)$, where $\hat{z}_i$ denotes the elevation of that surface point $\hat{p}_i = (x_i, y_i, \hat{z}_i)$ which has the same footprint (x_i, y_i) as the data point p_i. The point $\hat{p}_i$ is uniquely determined, provided the surface is elevated. The point $\hat{p}_i$, however, does not exist if the footprint (x_i, y_i) of the data point lies outside the footprint region of the surface. This fact gives rise to a natural concept of "overlap". Indeed, measures-of-fit should apply only to areas of overlap between the point cloud and the surface.

"Euclidean distance" is the conventional distance between point $p_i = (x_i, y_i, z_i)$ and the surface, that is, the smallest distance between p_i and any point on the surface. This distance is always defined. It is therefore necessary to exclude non-overlap areas. Roughly speaking, "Ray-directed distance" is defined as the distance measured in the direction of the laser beam. The overlap area is determined by those rays which meet both point cloud and surface.

In this report, only vertical distances are considered together with the RMS and ASD norms, respectively. In other words, the following measures of fit are considered:

- Vertical distance with the ASD norm
- Vertical distance with the RMS norm

Vertical distance is a natural choice for 2.5D, that is, TIN surfaces such as terrain representations.

3. OPTIMIZATION PROCEDURE

The intent of optimization-based registration methods is to identify a rigid transformation which -- when applied to a given point cloud – repositions the point cloud so as to minimize deviation from the given surface as quantified by a measure-of-fit. Rigid transformations may be characterized by a translation with parameters,

$$\Delta x, \Delta y, \Delta z,$$

combined with three rotations,

$$\Delta\varphi \,, \Delta\varepsilon, \Delta\theta,$$

the yaw, roll, and pitch, respectively [4]. The value of the measure-of-fit after transformation may be therefore considered as a function, F, of the above six transformation parameters, and this function is to be minimized.

In the case of the measures-of-fit based on vertical distance, this minimization met with several difficulties. First, in the neighborhood of an optimal parameter choice, the function F assumes the shape of a plateau with many local minima in close proximity of each other, and the quality of the registration tends to be sensitive to the choice of one of those local minima. Moreover, a global minimum is typically a disastrous choice. The automated procedure used for this work searches among neighboring local minima within a fixed radius and terminates if no improved measure-of-fit can be found. This procedure is, of course, very sensitive to the choice of starting point: If the starting point fails to be reasonably close to an acceptable "solution", the procedure may, in fact, lead to progressively worse registrations, all the while "improving" the measure-of-fit. The challenge is to find measures-of-fit which track the quality of registration. Those considerations lead to the following optimization process.

Having evaluated a particular set of parameters $(\Delta x^0, \Delta y^0, \Delta z^0, \Delta\varphi^0, \Delta\varepsilon^0, \Delta\theta^0)$, the translation $(\Delta x, \Delta y, \Delta z)$ is optimized, keeping the angle parameters fixed. Here, the two planar parameters $(\Delta x, \Delta y)$ are considered first. This planar optimization procedure will be described below in more detail. Once that optimization step is completed, the optimal vertical translation Δz may be determined in closed form by subtracting from the previous value either the mean or the median of all residuals, depending on whether the measure-of-fit is RMS or ASD based, respectively. If the new vertical translation parameter Δz changes from its previous value, then the planar optimization procedure resumes. Otherwise, the translation parameters $(\Delta x, \Delta y, \Delta z)$, are considered optimized, and the process moves to the optimization of the angle parameters.

In order to avoid minima that correspond to unacceptable registration results, the optimization of the planar parameters $(\Delta x, \Delta y)$ follows a "limited-horizon" search principle: Consider a circle of radius r around a current "point" $(\Delta x^0, \Delta y^0)$. Search for a "better" point $(\Delta x, \Delta y)$ within this circle only. If found, it becomes the center of another search within a radius of r around it. If no such improvement is found, then the planar search terminates and, as described above, the optimal vertical parameter Δz is determined. The procedure consistently decreases with measure-of-fit function F, and terminates with a particular "local" minimum.

For the experiments reported here a search radius, $r = 5$ cm, was used. If r is too large, it may lead the search astray. If r is too small, the search may terminate before reaching a "good" local minimum. As for most registration methods, the success of the search depends heavily on the choice of the initial parameters ("warm start" vs. "cold start").

There are, of course, many approaches to optimizing a function in a given region. Most advanced optimization algorithms for minimizing a function $f(\xi_1, \xi_2, \ldots, \xi_n)$, however, stipulate that a point $(\xi_1^0, \xi_2^0, \ldots, \xi_n^0)$ is a local minimum if perturbing single variables ξ_i^0 does not lead to lower values of f. This is not true for functions that are not continuously differentiable such as the measure-of-fit function F.

For this reason, we adopted a search based simply on sampling the search area. With $(\Delta x^0, \Delta y^0)$ as the center point, concentric sets of 20 trial points at equal angle increments of 22.5° are considered for 15 radii ranging from 0.02 cm to 5.0 cm. Generally, the search starts with radius 1.0 cm and if necessary, moves to successively smaller ones of the proposed radii. If reduction to the smallest radii fails to yield an improvement, then the search continues with increasing radii up to the search radius of 5 cm. As F is calculated for up to 15 x 20 = 300 parameter settings for each planar search step, the search procedure is computationally expensive and in this form, too slow for real-time purposes.

It was assumed, that the optimization of the angle parameters does not need to be similarly restricted except for reasons of efficiency. The planar rotation $\Delta\varphi$ is varied in single steps of 0.05° through a specified bracket and, for each value, the full translation $(\Delta x, \Delta y, \Delta z)$ is re-optimized. The selection of the roll and pitch parameters $\Delta\varepsilon$, $\Delta\theta$ has as yet not been automated, because previous experiments [4] suggested that -- for reasons as yet unclear -- minimizing those parameters actually worsened the quality of the registration.

Given a TINsurface, each specification of a horizontal "cut plane", at elevation z, defines both a "cut volume" and a "fill volume". The cut volume is bounded below by the cut plane and above by the surface as far as it extends above the cut plane. The fill volume is bounded above by the cut plane and below by the surface.

In the experiments to be described, evaluating a sequence of cut volumes was used for determining volume and height of the box once its surface and that of the surrounding floor has been modeled. More precisely, the idea is to create a sequence of cut volumes, starting with cut planes below the floor level, as indicated by a fill volume of 0, and proceeding by increasing the cut level z, in equal increments of 0.2 cm, until the cut plane clears the top of the box, as indicated by a cut volume of 0. As long as the cut plane remains below the surface, the cut volume will decrease linearly, the slope given by the area of the footprint region of the surface. As the cut plane starts to intersect the surface, a transition occurs to the regime in which the cut plane only intersects the box, and where the decrease is approximately linear with the box footprint as its slope. A similar transition is observed as the cut plane moves beyond the top of the box. The first transition marks the floor level. The second transition marks the top level of the box. The two transition elevations show up as "spikes" in the sequence of second differences [4]. The locations of these spikes define the elevations for floor and box top, respectively. Thus

box height = box top elevation - floor elevation
box vol. = cut vol. @ floor - cut vol. @ box top

4. EXPERIMENTS

As mentioned previously, four scans of a box were obtained. These scans were obtained with the scanner located "in front" of each box corner, i.e., in each scan, the box top and two sides of the box were visible. The scans are clockwise labeled C, D, E, F. The four scans were visually transformed so that they were in rough alignment to each other, Figure 1. This was necessary as the registration program assumes small transformations – this is true for most registration programs.

Two series of experiments were conducted: 1) registration against an exact model of the box and 2) registration of a point cloud to another point cloud. For the second series, point cloud C was arbitrarily selected as the reference point cloud. Parameters for both series include: 1) starting point of registration, cold vs. warm start and 2) measure-of-fit, ASD or RMS. For the starting point of registration, a cold start was a start at the origin where $(\Delta x, \Delta y, \Delta z) = (0, 0, 0)$ and a warm start was a start off the origin at some point $(\Delta x, \Delta y, \Delta z)$ close to the expected final registration point [4]. For the cold start, the φ value ranged from $-2.5°$ to $2.5°$, and the increment size was $0.05°$. For the warm start, the range of the φ value was $1°$ starting close to the final value [4]. The values for ε and θ were set to zero and not changed.

The results of the automated registration are given in Tables 1 and 2. Summary of results:

- ASD vs. RMS: ASD yields a better registration in terms of volume and box height
- Registration to box: ASD volume error on the order of 2 % (cold and warm starts), RMS volume error of 12 % (warm start) and 137 % (cold start) – suspect registration got stuck on local minima.
- Point cloud to point cloud registration: ASD volume error on the order of 10% (cold and warm starts), RMS volume error 13 % (warm start) and 36 % (cold start).
- Warm vs. cold start: For RMS, significant reduction of volume error for warm start. For ASD, no significant difference in volume error for warm vs. cold start.

Figures 2 (registration to box) and 3 (point cloud to point cloud registration) show the final registration of the four scans. The reason for the misalignment of the last scan in Figure 3 is being examined.

5. CONCLUSIONS

The experiments reported in this paper clearly show the importance of selecting the correct measure-of-fit for optimizing the registration. In these experiments which were based on vertical distance, the ASD norm produced a better registration than the RMS norm in terms of box volume and box height. Future work will include using other measures-of-fit such as the "ray-directed" distance.

6. REFERENCES

1. Besl, P., and McKay, N., "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, 1992, pp. 239-256.
2. Cheok, G. S. and Leigh, S., Calibration of a Laser Scanner, in preparation.
3. Rusinkiewisc, S. and Levoy, M., "Efficient Variants of the ICP Algorithm," in *Proc., International Conference on 3D Digital Imaging and Modeling* (3DIM2001)," IEEE Computer Society, 28 May – 1 June, 2001, pp. 145-152.
4. Witzgall, C. and Cheok, G. S., Registering 3D Point Clouds: An Experimental Evaluation, *NISTIR 6743*, NIST, Gaithersburg, MD, May, 2001, 41 pp.
5. Witzgall, C. and Cheok, G. S., Triangular Meshing for Volume Determination, in preparation.

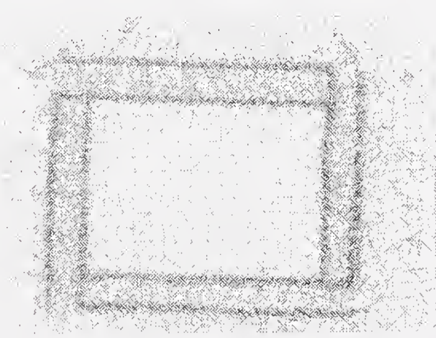


Figure 1. Triangulation of Footprint of Roughly Aligned Scans - Starting Point for Registration.

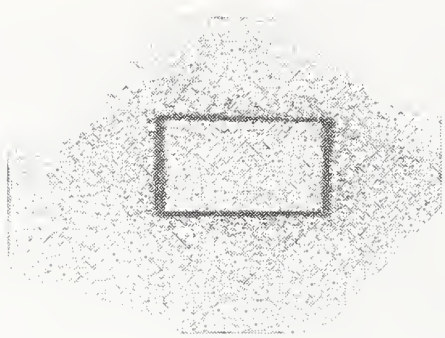


Figure 2. Registration to Box – Final Registration.

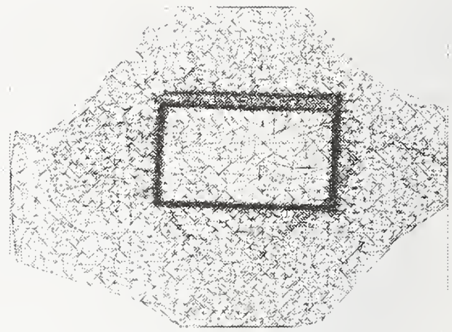


Figure 3. Point Cloud to Point Cloud Registration – Final Registration.

Table 1. Registration of Point Cloud to Exact Box Model.

Start	Norm Used		Final Registration Values: Δx (cm)	Δy (cm)	Δz (cm)	Δφ (°)	Δε (°)	Δθ (°)	ASD (cm)	RMS (cm)	Volume (m^3)	Volume Error (%)	Box Height (cm)	Box Ht. Error (%)
Start at (Δx, Δy, Δ z) = (0, 0, 0) (Cold Start)	ASD	C	6.550	-8.790	-1.000	2.45	0	0	19.101	33.877	1.727	1.63	92.0	0.61
		D	6.605	8.069	0.100	2.35	0	0	25.206	38.890				
		E	14.052	0.443	-3.900	-0.45	0	0	24.910	38.898				
		F	16.984	13.718	-1.900	1.20	0	0	24.089	38.111				
	RMS	C	0.370	-6.318	15.072	2.00	0	0	24.071	30.557	4.026	136.97	92.6	1.27
		D	1.956	2.566	-27.112	2.50	0	0	27.395	31.056				
		E	18.636	-13.208	18.423	2.30	0	0	27.565	33.550				
		F	19.047	12.148	-26.209	2.40	0	0	27.152	30.860				
Start off origin (warm start)	ASD	C	7.443	-9.829	-0.800	2.80	0	0	19.044	33.770	1.671	-1.65	92.8	1.49
		D	6.048	8.564	0.100	2.20	0	0	25.211	38.895				
		E	20.246	-12.027	-1.900	2.25	0	0	25.005	38.838				
		F	25.812	6.232	-0.300	2.95	0	0	24.349	38.335				
	RMS	C	0.372	-6.308	15.072	2.00	0	0	24.071	30.557	1.906	12.20	97.5	6.63
		D	8.936	8.038	20.235	2.80	0	0	27.540	33.569				
		E	18.626	-13.222	18.423	2.30	0	0	27.565	33.550				
		F	26.623	6.326	19.998	2.65	0	0	27.080	32.873				

Table 2. Registration of Point Cloud to Point Cloud.

Start	Norm Used		Final Registration Values: Δx (cm)	Δy (cm)	Δz (cm)	Δφ (°)	Δε (°)	Δθ (°)	ASD (cm)	RMS (cm)	Volume (m^3)	Volume Error (%)	Box Height (cm)	Box Ht. Error (%)
Start at (Δx, Δy, Δ z) = (0, 0, 0) (Cold Start)	ASD	C	0	0	0	0	0	0	NA	NA	1.861	9.53	93.3	2.03
		D	-7.889	17.475	-0.256	0.35	0	0	16.910	25.728				
		E	20.959	12.599	-1.456	0.95	0	0	14.020	25.250				
		F	12.406	18.133	0.504	0.60	0	0	20.044	32.902				
	RMS	C	0	0	0	0	0	0	NA	NA	2.304	35.59	92.6	1.27
		D	-8.122	17.584	-2.904	0.25	0	0	17.684	25.629				
		E	23.224	12.994	-7.367	1.30	0	0	16.829	25.761				
		F	14.741	31.937	-7.049	1.20	0	0	18.466	26.677				
Start off origin (warm start)	ASD	C	0	0	0	0	0	0	NA	NA	1.859	9.42	93.0	1.71
		D	-7.829	17.43	-0.267	0.35	0	0	16.908	25.727				
		E	17.587	17.063	-1.521	0.10	0	0	14.123	25.330				
		F	10.21	22.372	0.508	-0.10	0	0	20.887	33.774				
	RMS	C	0	0	0	0	0	0	NA	NA	1.913	12.57	93.6	2.36
		D	-8.129	17.6	-2.839	0.25	0	0	17.661	25.629				
		E	16.261	3.19	1.348	0.05	0	0	17.335	27.595				
		F	12.499	22.903	1.773	0.10	0	0	17.140	25.873				

Human-Assisted Object Fitting to Sparse Range Point Clouds for Rapid Workspace Modeling in Construction Automation

by

Soon-Wook Kwon[1], Katherine A. Liapi[2], Carl T. Haas[3], S.V. Sreenivasan[4], and Jared T. McLaughlin[5]

ABSTRACT: In large-scale construction sites there are constant needs for rapid recognition and accurate measurement of objects so that on-site decisions can be made quickly and safely. Current methods involve full area laser range scanning systems that can produce very detailed models of a scanned scene, however the computational and data acquisition time that is required precludes the methods from being used for real time decision making. This paper presents algorithms to fit objects to sparse point clouds of measured data in a construction scene, that significantly decrease data acquisition time, and computational and modeling time. Two basic fitting and matching algorithms that address construction site material of cuboid and cylindrical shapes are discussed. Experimental results that indicate that the proposed algorithms assist an operator to create models of construction objects rapidly and with sufficient accuracy are also presented.

KEYWORD: CONSTRUCTION AUTOMATION; LASER RANGE FINDER; LEAST SQUARES METHOD; OBJECT FITTING; OBJECT MATCHING; WORKSPACE MODELING

1. INTRODUCTION

Using automated or semi-automated equipment on a large construction site requires rapid recognition and accurate measurement of objects in the workspace so that timely on-site decisions can be made. Most methods for modeling work environments rely on analyzing dense point cloud data, which requires computationally intensive processing, and usually takes much longer than the ongoing construction operation. Low accuracy in extracting objects from dense clouds is an additional limitation of full range scanning methods. Since most objects in a construction site are known and man made, they can be graphically generated and stored in object database as parametrically defined object classes [1]. By exploiting a human operator's ability to recognize objects in a construction scene, pre-stored graphic representations of construction objects can be matched and fitted to sensed data from 3D position sensors deployed in the construction environment [2][3].

The ability to extract models of real

[1] Graduate Research Assistant, Dept. of Civil Eng., The University of Texas, Austin, TX 78712, swkwon@mail.utexas.edu
[2] Assist. Professor, Dept. of Civil Eng., The University of Texas, Austin, TX 78712, kliapi@mail.utexas.edu
[3] Professor, Dept. of Civil Eng., The University of Texas, Austin, TX 78712, haas@mail.utexas.edu
[4] Assoc. Professor, Dept. of Mech. Eng., The University of Texas, Austin, TX 78712, sv.sreeni@mail.utexas.edu
[5] Engineer, Los Alamos National Lab, NM, jtmclau@sandia.gov

world objects in a construction workspace for equipment operations from only a limited number of scanned points is a significant advantage of this approach over full range scanning methods that require intensive computational load because of range data processing for dense point clouds which consists of tens of thousand data points.

This paper presents algorithms that accurately fit and match objects, with regard to location and orientation, to sparse point clouds which have less than 50 scanned points for each object in a construction scene. The implementation of the algorithms will allow a human operator to rapidly construct a world model from unfiltered real-world range data.

With respect to the geometric primitives most frequently encountered in a construction site, it appears that a few types of objects can be used to model a wide range of construction scenes [4]. Cuboids can be used for fitting and matching structural objects such as columns, box-beams and walls and finishing objects. Cylinders can be used to fit and match chemical pipes, ventilation pipes, and concrete piles. The following fitting and matching algorithms were developed for each primitive:

1. Cuboid algorithm
2. Cylindrical object algorithm

2. EXPERIMENTAL SETUP AND HUMAN ASSISTED OBJECT FITTING AND MATCHING PROCESS

A single-axis laser range finder, a pan and tilt unit, and a personal computer were used for the experimental set up (Figure 1). The single-axis laser range finder (DistoMemo) that is mounted on the pan and tilt unit is designed not only for hand-held operation, but also for computer use through interface. The measurements can be remotely taken and transferred directly into the computer. The range of measurement of the laser range finder is 100 m with accuracy of ±3 mm. The step size of the tele-operated pan and tilt unit, which controls the laser range finder, is of high resolution (0.0128571°/step) and its maximum speed is a little over 60°/second. Its error is 0.2 cm for every 10 m in motion.

The sparse points cloud is acquired by operator picking points to each object using single-axis laser range finder. The modeling process involves the following functions:

1. Select object for scanning (by operator)
2. Acquire sparse point cloud data in the form of range images
3. Convert range data into xyz coordinates
4. Analyze the features of each surface of the object
5. Match all of the object surfaces with the model's surfaces using matching algorithms
6. Fit the object into the point cloud using fitting algorithm

Figure 2 displays a process diagram of these functions.

3. OBJECT FITTING AND MATCHING ALGORITHMS

Graphical workspace modeling can improve construction equipment control and operations. Equipment operators can use graphical workspace models as an interactive visual feedback tool during equipment controls [2][5].

For the rapid modeling of construction site objects from sparse point clouds two basic algorithms were developed that address construction site objects of cuboid and cylindrical shape. Since these two types of primitives consist of 6 planar surfaces (cuboid), and two planar surfaces and one curved surface (cylinder), the algorithms were developed as a surface based fitting and matching method. Algorithm development and revisions were based on lab experiments.

By using these algorithms we achieve: (1) accurate and reliable methods to save computational cost and time, (2) improved fitting and matching methods to attain real-time execution, and (3) increased modeling accuracy with operator's assistance.

The following sections explain the fitting and matching methods which were developed and used for rapid workspace modeling:

3.1 Cuboid Algorithm

This section describes how to fit a sparse points cloud to a cuboid's surfaces using the k-nearest neighbors and the least squares methods. There is an assumption of this process that three surfaces of cuboid should be visible in order to acquire data points.

3.1.1 Point Segmentation Using K-Nearest Neighbors Method

To find the nearest points for all measured points on a cuboid, a k-nearest neighbors algorithm was used. The algorithm finds the nearest two points by computing all the distances from a scanned point to all other points [6]. After determining two nearest neighbors for each scanned point, a group of three-point sets was found. Then, a normal vector for each three-point set was computed. By analyzing normal vectors, the scanned points were segmented by each cuboid surface.

3.1.2 Plane Optimization Using the Least Squares Fitting Method

The least squares method [7] was used for the best-planar fit of point sets on each surface of the cuboid after segmentation was applied.

Since in a planar regression, Y is to be regressed on two independent variables X and Z, a relationship, where both X and Z, are calculated as deviations from their means, was used:

$$E(Y_i) = \alpha + \beta_i \cdot X_i + \gamma \cdot Z_i \tag{1}$$

For any given combination of X_i and Z_i the expected yield $E(Y_i)$ is a point directly above the plane, shown as a hollow dot in Figure 3. The actual value of the component Y_i of an observed point is somewhat greater than its expected value and is shown as a solid dot lying on the plane. The difference between the observed and expected values of Y_i is shown by the error term e_i and thus the observed value Y_i is expressed as its expected value plus the error term e_i:

$$Y_i = \alpha + \beta_i \cdot X_i + \gamma \cdot Z - e_i \tag{2}$$

While moving along the x-direction, β_i is interpreted as the slope of the plane. In the same way γ is the subsidiary effect of z. To minimize the error sum of the squares a coefficient is used:

$$\sum_{i=1}^{N}(Y_{i\_} - \hat{Y})^2 = \sum_{i=1}^{N}(Y_i - \hat{\alpha} + \hat{\beta} \cdot X_i + \hat{\gamma} \cdot Z_i)^2 \tag{3}$$

Taking the partial derivatives of the above expression with respect to $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$, and setting them to zero, finally alpha, beta, gamma are found. Using this expression, the three optimized surfaces of the cuboid are computed (Figure 3).

After segmenting all scanned points by the three surfaces of the cuboid, the points were projected onto the optimized surface to compute dimensions.

3.1.3 Determining Intersecting Edges and Computing Dimensions

The three surface planes of the cuboid, from which range data were received, intersect at a point, and each two planes intersect at a line. The intersection of the two planes of the cuboid was found by solving the two linear equations representing the planes. After applying this for all three surfaces of the cuboid, the three edges of the cuboid were determined and matched. A vertex of the cuboid was also determined. Figures 4 and 5 show the results of point segmentation, and matching vertex process.

Once the three edges of the cuboid were defined, the dimensions of the cuboid were determined as follows: By computing the distances of all measured points on each surface to each one of the already defined edges of the same surface, the furthest point from each edge was found. The distances of the furthest points on the surfaces from the three intersecting edges represent the dimensions of the cuboid.

Figure 6 shows a fitted and matched model of an object after the application of the cuboid algorithm.

3.2 Cylindrical Object Algorithm

Four parameters are required for fitting and matching a solid cylinder: a scalar radius r; an axis vector, a; a center point to determine the axis vector, $c = (X_c, Y_c, Z_c)$ and a set of scanned points $g = \{(X_i, Y_i, Z_i)\}$ to find out the boundary of the cylinder. To determine the normal vector, the "k-nearest neighbors method" was used. Then, by analyzing normals, the scanned points were segmented by surface (planar or curved). Subsequently, by projecting all points on the curved surface onto the planar surface, parameters r and c were estimated. The least squares method was also used to optimize the curved surface. The radius of the cylinder was found as the distance from the center of the circle to any point on the optimized curve. Projected points on the planar surface are considered end points of different chords in the circle and used to estimate its center $\hat{c}$. An initial estimate of the radius, $\hat{r}$, is found by $\hat{r} = mean(|\hat{c} - k'|)$ (k' = {the points on the optimized curve of planar surface}). Then the final values of a, c, and r are found by applying the least squares method to all scanned data (Figure 7).

4. EXPERIMENTAL RESULTS AND CONCLUSIONS

The fitting and matching algorithms discussed in this paper, are an integral part of a method that involves several other functions such as: human object recognition, collecting of range information, grouping of scanned points, and computing dimensions to final fitting and matching. A basic feature of the method is that it takes advantage of human cognitive ability to recognize and classify objects in the workspace; that is a human operator initiates scanning, recognizes objects, and controls the system for data acquisition. In addition fitted and matched objects are verified by the operator and then inserted into the workspace model.

Experiments were conducted to determine the efficiency of the human assisted modeling method. The algorithms, which are based on the least squares method, were found to be useful for modeling construction objects of cylindrical and cuboid shapes. They were applied to determine the width, depth, and height of cuboids, and the diameter, and height of solid cylinders including the location and orientation.

Table 1 shows an example of experimental results of a cuboid fitting and matching process. The test results of the algorithms present approximately less than 1-degree angular deviation between model and real objects' axis. Respectively in all tests the size difference between the modeled and the actual object's surfaces is less than 5 %. For increased accuracy further modifications of the algorithms are required. In general low deviation values and the low modeling times in Table 1 indicate that a system based on the above geometric algorithms and a human-guided simple laser range finder can model construction objects rapidly and with sufficient accuracy.

5. REFERENCES

[1] Cho, Y., Haas, C., Saidi, K., Liapi, K., and Sreenivasan, S.V.(2001), "Rapid Local Area Modeling for Construction Automation" Proc., 18th International Symposium on Automation and Robotics in Construction, Krakow, September

[2] Kim, Y.S. and Haas, C. (2000) "A Model for Automation of Infrastructure Maintenance using Representational Forms", Journal of Automation in Construction, Vol.10, No.1.

[3] Cho, Y. (2000), "Human-Assisted Rapid Workspace Modeling for Construction Equipment Operations," PhD Dissertation, The University of Texas at Austin

[4] Johnson A., Hoffman R., Osborn J., and Hebert M.(1997), "A System for semi-automatic Modeling of Complex Environment" Proc. IEEE International Conference on 3-D Digital Imaging and Modeling, Ottawa, May

[5] McLaughlin, J. T. (2002), "Rapid Human-Assisted Creation of Bounding Models for Obstacle Avoidance in Construction," Master's Thesis, The University of Texas at Austin

[6] Duda R. O., Hart P.E., and Stork D.G. (2001), "Pattern Classification" Wiley Inter-Science

[7] Duran B.S., and Amir-Moez A.R. (1973) "Linear Algebra of The Plane", Texas Tech University

ACKNOWLEDGEMENTS

This paper is based on the research funded by the National Science Foundation (Grant#: CMS-0000137) and the National Institute Standard and Technology (Project NBR: NA1341-02-W-0742). The authors gratefully acknowledge their financial support and encouragement throughout this study.

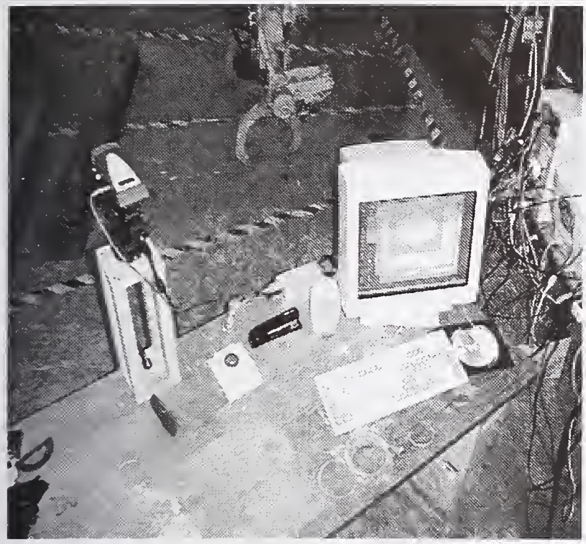


Figure 1. Experimental Setup

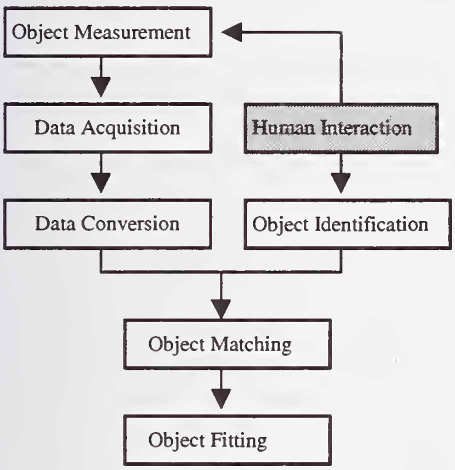


Figure 2. Fitting and Matching Process

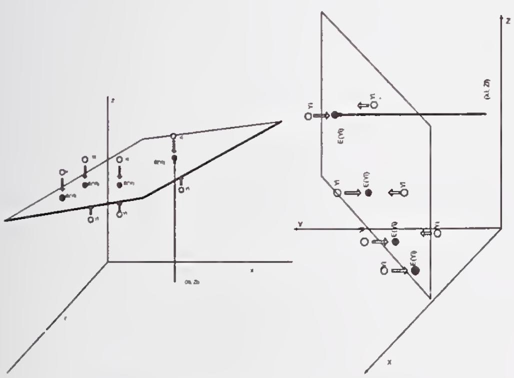


Figure 3. Surface Optimization

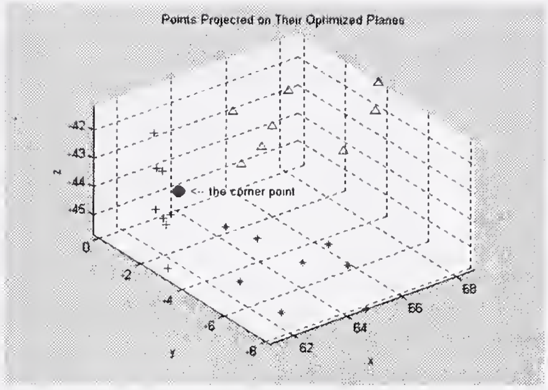


Figure 4. Matching Points and Segmentation

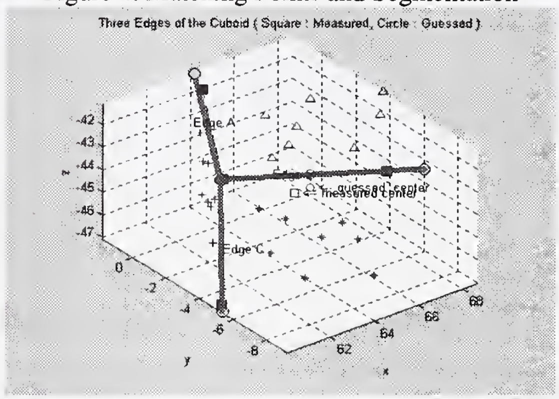


Figure 5. Three Edges of a Cuboid and its Centroid

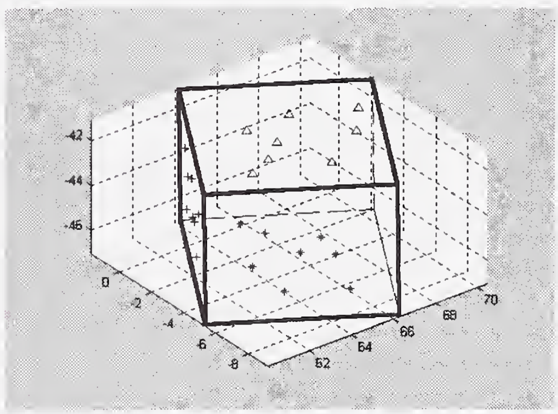


Figure 6. Fitted and Matched Cuboid

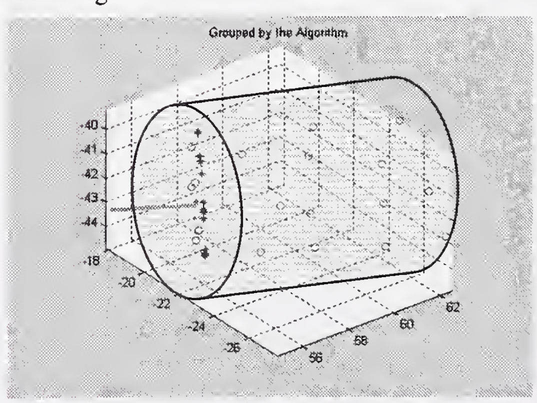


Figure 7. Fitted and Matched Cylinder

Table 1. Test Results of the Cuboid Algorithm

Matching Point (vertex)		Modeled Object	Actual Object	Deviation
Object 1	x	57.749	57.816	0.067
	y	11.697	11.733	0.037
	z	-35.447	-35.464	-0.017
Object 2	x	59.981	60.037	0.056
	y	-5.203	-5.136	0.067
	z	-41.370	-41.543	-0.173
Object 3	x	59.967	60.028	0.061
	y	-5.182	-5.108	0.073
	z	-41.354	-41.492	-0.138
Object 4	x	59.918	60.032	0.114
	y	-5.170	-5.109	0.061
	z	-41.210	-41.621	-0.411
Angular deviation between edges	Edge A			1.089
	Edge B			1.824
	Edge C			0.927
Measuring + Computing time	pts	Measuring (sec.)	Computing (sec.)	
Object 1	16	30.00	5.87	35.87
Object 2	18	50.00	5.66	55.66
Object 3	22	60.00	6.48	66.48
Object 4	26	80.00	5.60	85.60

Reconstructing Images of Bar Codes for Construction Site Object Recognition[1]

by

David E. Gilsinn[2], Geraldine S. Cheok[3], Dianne P. O'Leary[4]

ABSTRACT: This paper discusses a general approach to reconstructing ground truth intensity images of bar codes that have been distorted by LADAR optics. The first part of this paper describes the experimental data collection of several bar code images along with experimental estimates of the LADAR beam size and configuration at various distances from the source. Mathematical models of the beam size and configuration were developed and were applied through a convolution process to a simulated set of bar code images similar to the original experiment. This was done in order to estimate beam spread models (beam spread models are unique to each specific LADAR) to be used in a deconvolution process to reconstruct the original bar code images from the distorted images. In the convolution process a distorted image in vector form g is associated with a ground truth image f and each element of g is computed as a weighted average of neighboring elements of f to that associated element. The deconvolution process involves a least squares procedure that approximately solves a matrix equation of the form $Hf = g$ where H is a large sparse matrix that is made up of elements from the beam spread function.

KEYWORDS: bar codes, deconvolution, image processing, LADAR, object recognition, sparse matrix.

1. INTRODUCTION

Imaging sensors such as LADARs (laser distance and ranging devices) are used to rapidly acquire data of a scene to generate 3D models. They are used to obtain two- or three-dimensional arrays of values such as range, intensity, or other characteristics of a scene. Currently available LADARs can gather four pieces of information – range to an object; two spatial angular measurements; and the strength of the returned signal (intensity). Various methods are used to convert the data, which are collected in the form of point clouds, into meaningful 3-D models of the actual environment for visualization and scene interpretation. The points within a point cloud are indistinguishable from each other with regard to their origin; i.e., there is no way to tell if a point is reflected from a tree or from a building. As a result, the methods used to generate the models treat all points identically and the results are indistinguishable "humps/bumps" in the scene. Current surface generation methods using LADAR data require intensive manual intervention to recognize, replace, and/or remove objects within a scene. As a result, aids to object identification have been recognized by the end users as a highly desirable feature and a high priority area of research.

The use of bar codes or UPC (Universal Product Code) symbols has become the universal method for the rapid identification of objects ranging from produce to airplane parts. The same method could also be used to identify objects within a construction scene. This would involve

---

[1] Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States

[2]National Institute of Standards and Technology (NIST), Information Technology Lab, Mathematical and Computational Sciences Division, Gaithersburg, MD 20899-8910: dgilsinn@nist.gov

[3]NIST, Building and Fire Research Lab, Construction Metrology and Automation Group, Gaithersburg, MD, 20899-8611: cheok@nist.gov

[4] Computer Science Department, University of Maryland, College Park, MD 20742: oleary@nist.gov

using the LADAR to "read" a bar code. The concept is to use the intensity data from the LADAR to distinguish the bar pattern. The advantage of this concept is that no additional hardware or other sensor data is required to gather the additional data. The basis for this concept lies in the high intensity values obtained from highly reflective materials.

The challenges are the ability to read the bar code from 100 m or greater, distinguish bar code points from the other points in the scene, capture a sufficient number of points to define the bar code for correct identification, and the ability to read bar codes that are skewed.

At ISARC 2001 [1] methods of determining the appropriate material for the bar codes and determining if the use of bar codes was viable were reported. The results from these efforts showed that at distances beyond 20 m, the intensity images were too blurred to be readable and that image processing techniques were necessary to reconstruct the image. The blurring or convolution of the image is a result of the low resolution (number of pixels/unit area; a consequence of the instrument's angular resolution) of the intensity images at larger distances and of distortion of the intensity image by the LADAR optics. As a result, an investigation of possible methods to de-blur (deconvolve) the intensity images was conducted. Deconvolution of the image involves reversing the convolution implying that if the convolution process was known, the image may be reconstructed. The results of the deconvolution study are reported here.

2. BEAM SIZE AND DIVERGENCE ESTIMATION

The data for determining the beam size as a function of distance was obtained as part of an experiment to determine the range accuracy of the LADAR as a function of the angle of incidence of the laser beam and distance. An infrared viewer was used to see the projection of the laser on the target so that an outline of the beam could be drawn. The outline of the beam was drawn by two or more observers and the measurements were averaged. The LADAR used consisted of three laser diodes. The projection of the laser beam on the target was seen as a bright rectangle for distances less than 10 m and three bright vertical bands separated by dark bands for distances greater than 10 m. A schematic of the beam is shown in Figure 1. The measurements are given in Table 1.

The average vertical beam divergence was 2.14 mrad (σ = 0.39 mrad) and the average horizontal beam divergence, excluding an outlier (negative divergence) was 1.86 mrad (σ = 0.44 mrad). The average beam divergence (horizontal and vertical combined) was 2.01 mrad (σ = 0.43 mrad) - compared with the manufacturer's specified divergence of 3 mrad. The lower experimental value was likely a result of the inability of the unaided human eye to detect the faint edges of the laser beam projection. A schematic of the diverging beam is shown in Figure 2.

3. OPTICS MODEL FOR BEAM DISTORTION

Discretize the ground truth image by defining: $0 = y_1 < y_2 < \ldots < y_{nf} = 1$, $0 = z_1 < z_2 < \ldots < z_{nf} = 1$, $\Delta y_i = y_{i+1} - y_i, = \Delta z_j = z_{j+1} - z_j. = \Delta$. The intensity at patch $\Delta y_i \Delta z_j$ is given by $f(y_i^*, z_j^*)\Delta y_i \Delta z_j$ where f is a function of (y, z) expressing the intensity response at some point (y_i^*, z_j^*) in the patch. Due to distortions, the LADAR image of the response from the bar code surface is smeared out into some form of blurred spot shown in Figure 3. The distorted image will be taken as a subset (to be defined below) of the ground truth image for simulation purpose. Points in the distorted image will be identified by (Y, Z) and those in the ground truth image by (y, z). These are different notations for points in the same axis system. The distortion at a point (Y, Z) due to a point (y, z) can be described by a function $h(y, z; Y, Z)$, called the Beam Spread Function. For most practical purposes, the Beam Spread Function can be considered translation invariant in the sense that its distortion value only depends on the distance between (Y, Z) and (y, z) so that h has the form

$h(Y\text{-}y,\ Z\text{-}z)$ (see Figure 4). The incremental distortion effect at (Y, Z) due to a neighboring patch of (y, z) is $g(Y, Z) = h(Y\text{-}y, Z\text{-}z) f(y, z)\Delta y \Delta z$. To describe the total effect $g(Y, Z)$ of all of the points (y, z) in the ground truth image one sums over all of the patches in the ground truth image. As the number of grid points nf in the ground truth image grows and the patch size tends to 0, the sum can be replaced by an integral. This integral is called a convolution integral.

$$g(Y,Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(Y-y, Z-z) f(y,z)\,dy\,dz$$

Assume that the distorted image g is sampled at points $\left(Y_p = (p-1)\Delta, Z_q = (q-1)\Delta\right)$ for p, q = 1, ... , ng, forming a vector G of ng^2 elements. Let h be sampled at $(\mathrm{ma})^2$ points $(i\Delta, j\Delta)$ for i, j = $-k$, $-k$+1, ..., k-1, k. This gives a total of ma^2 data values $\hat{H}$, where $ma = 2k+1$, and we define all other values of our approximation of h to be zero. Assume that we approximate the ground truth image by $f((i-1)\Delta, (j-1)\Delta) \approx F(i,j)$ for i, j = 1, ... , nf, and again arrange these values as a vector. Then our convolution integral can be approximated by

$$G(p,q) \approx \sum_{i=1}^{nf}\sum_{j=1}^{nf} \hat{H}(p-i, q-j) F(i,j) \Delta y_i \Delta z_j$$

This gives us ng^2 linear conditions on nf^2 unknowns $F(i, j)$. Let H be the matrix derived from the values $\hat{H}$ so that our approximation becomes $G \approx HF$.

4. COMPUTATIONAL ASPECTS FOR IMAGE RESTORATION

Since the ground truth image F is larger than the distorted image G there are more degrees-of-freedom involved in reconstructing F from a measured G. A computable approach is to determine F in a least squares manner to satisfy

$$\min_F \|HF - G\|.$$

This is an ill-posed problem since there may not be a single solution. In fact it is known that if F is combined with high frequency sinusoidal data, then, under convolution, it produces the same G. A penalty term can be added to this minimization problem that puts a premium on the size of F selected. Introduce $\lambda > 0$ and form the following minimization problem

$$\min_F \left\{ \|HF - G\| + \lambda \|F\| \right\}$$

The second term is called a regularization term and its function is to control the magnitude of the final F. In practice, λ is selected as a small positive number.

We solve the least squares problem using the LSQR algorithm of Paige and Saunders [2]. This algorithm never modifies elements of H and uses the matrix only to form products Hy with various vectors y. Thus we need only store the ma^2 coefficients of $\hat{H}$ rather than the ng^2 x nf^2 matrix H.

5. CONCLUSIONS

In order to determine the effect of beam spread models on ground truth images, one has to determine the nature of ground truth. The LADAR returns intensity levels in the range 0 to 255. From measured images, it was determined that the intensity of the return signal of the board on which the bars were mounted was approximately 150 and the intensity levels of the bars were approximately 250. Simulated ground truth data files were created for three sets of barcodes (see ISARC 2001[1]) with an example of 25.4 mm (1 in) bars shown in Figure 5. Based upon the measurements of the beam spread function three beam matrices were created to represent the spread function at 10 m, 20 m, and 40 m. Since it was difficult to obtain a precise measurement of the beam spread function, matrices representing the three spatial beam spread configurations were created. They were defined in such a manner that the area representing the dark regions was set to zero and the constant value assigned to the light areas was chosen so that the volume under the bright bars summed to unity. With the simulated barcodes and the simulated beam spread functions given, convolution calculations were performed in order to determine how close the simulated distorted images compared to the measured images. The simulated distorted images did not produce the horizontal spread distortion that the measured images produce. Although both the ground truth and beam spread images were simulated it is likely that the lack of prediction was produced by poorly understood beam spread functions. Since the preliminary measurements of the beam spread functions using an infrared scope were crude, this is not surprising. What was surprising, though, was that when the beam spread function at 10 m was used to deconvolve two images it was possible in a limited fashion to recover the basic details of the ground truth images. In particular, for the case of 25.4 mm (1 in) bars at 10 m, the measured data is given in Figure 6 and shows the lower three bars enmeshed together. The simulated beam spread function for 10 m was used to recover the ground truth image, and the result is shown in Figure 7. An attempt was then made to reconstruct the ground truth image for the 50.8 mm (2 in) bars at 10 m with the same simulated beam spread function. Figure 8 shows that the full reconstruction was not completely obtained. This suggests that the beam spread function might be influenced by the individual image being deconvolved, especially in the presence of noise. It, therefore, is clear that the nature of the beam spread function and its relation to the image being deconvolved is significant. An attempt was then made to construct the beam spread function using the least squares algorithm but in this case, set the matrix H to be the ground truth image and the unknown F to be the unknown beam spread matrix. Figure 9 shows the distorted image created for (25.4 mm (1 in) bars using the best fit Beam Spread matrix. It is very close to the actual data measured for the same bars as given in Figure 6. All of the results, though, point to the fact that reconstructing ground truth from distorted LADAR images is critically dependent on knowledge of the Beam Spread Function and how it relates to individual images.

6. REFERENCES

1. W. C. Stone, G. S. Cheok, K. M. Furlani, D. E. Gilsinn, "Object Identification Using Bar Codes Based on LADAR Intensity", *Proc. of the 18^{th} IAARC / CIB / IEEE / IFAC International Symposium on Automation and Robotics in Construction*, ISARC 2001, 10-12 September, 2001, Krakow, Poland.
2. C. C. Paige, M. A. Saunders, 'Algorithm 583: LSQR: Sparse Linear Equations and Least Squares Problems'. *ACM-Trans. Math. Software*, Vol. 8, No. 2, Jun. 1982, pp. 195-209
3. G. Golub, W. Kahan, 'Calculating the singular values and pseudo-inverse of a matrix', *J. SIAM Numer. Anal.*, Ser. B, Vol. 2, No. 2, 1965, pp. 205-224.

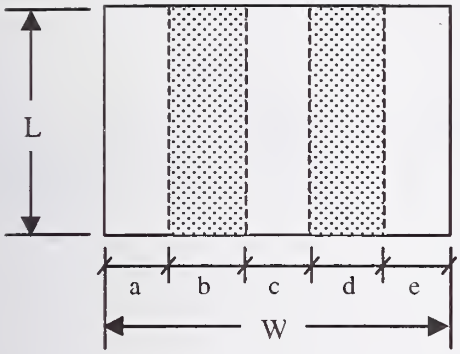


Figure 1. "a", "c" and "e" are the widths of the bright bands and "b" and "d" are the widths of the dark bands.

	10 m	20 m	40 m
L	27.5	56.2	108
W	52.4	64	111
a	-	12.7	15
b	-	13.2	25
c	-	14.7	22.3
d	-	11.7	30.5
e	-	12	18

Table 1. Average height and widths in millimeters of the bright and dark bands of the beam shown in Figure 1.

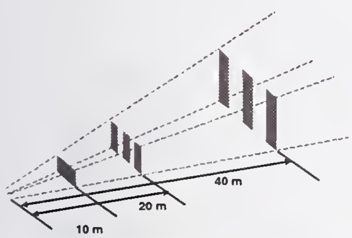


Figure 2. Beam divergence.

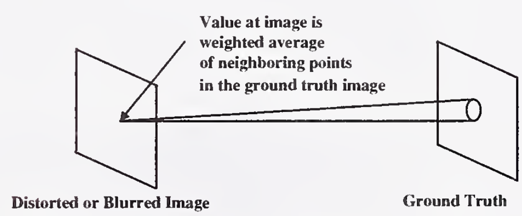


Figure 3. Distortion on the image caused by averaging of pixels in the ground truth image.

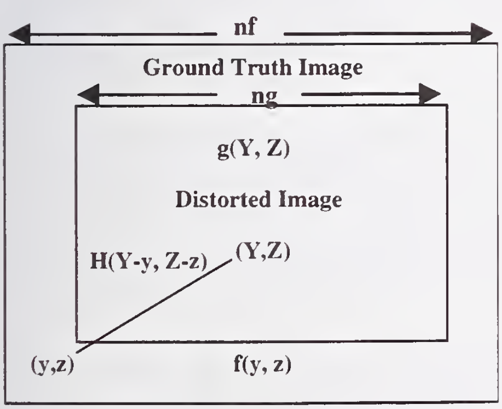


Figure 4. Distorted image overlaid on the ground truth image.

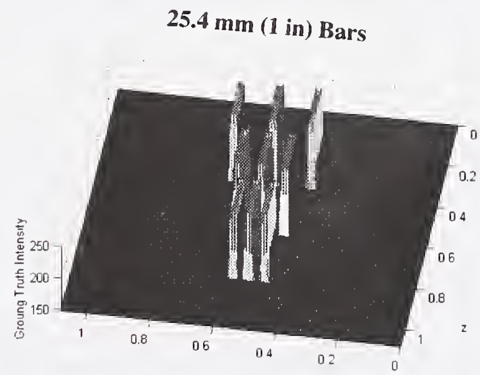


Figure 5. Simulated 25.4 mm (1 in) barcodes.

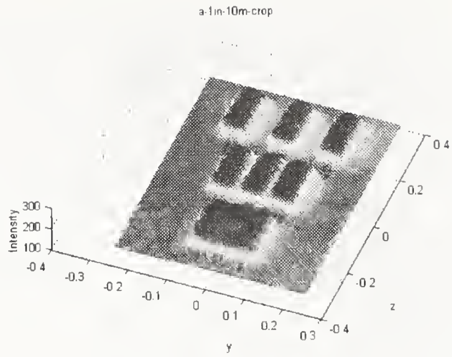


Figure 6. Measured intensity data by LADAR for 25.4 mm (1 in) bars.

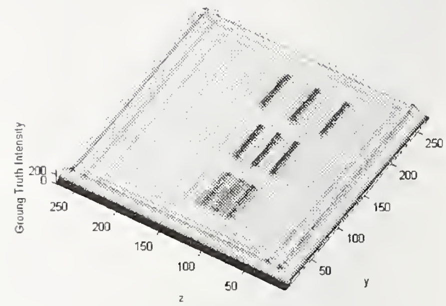


Figure 7. Reconstructed 25.4 mm (1 in) bars at 10 m.

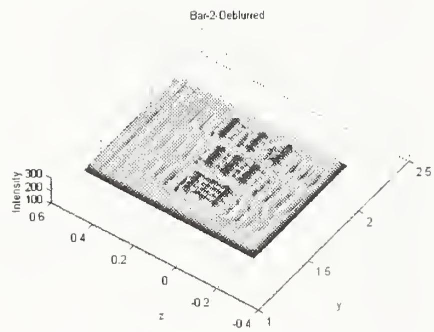


Figure 8. Partial reconstruction of 50.8 mm (2 in) bars at 10 m.

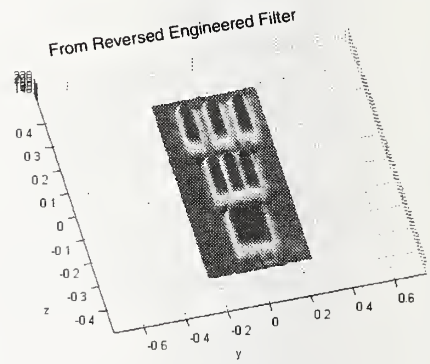


Figure 9. Distorted 25.4 mm (1 in) bars based on a least squares estimate of the Beam Spread Function.

ARTIFICIAL INTELLIGENCE BASED QUALITY CONTROL OF AGGREGATE PRODUCTION

by

Hyoungkwan Kim[1], Carl T. Haas[2], Alan F. Rauch[3]

ABSTRACT: This paper discusses a quality control method, based on artificial neural networks, that enables a plant operator to quickly detect property variations during the production of stone aggregates. The group texture concept in digital image analyses, two-dimensional wavelet transforms, and artificial neural networks are reviewed first. An artificial intelligence based aggregate classification system is then described. This system relies on three-dimensional aggregate particle surface data, acquired with a laser profiler, and conversion of this data into digital images. Two-dimensional wavelet transforms are applied to the images and used to extract important features that can help to differentiate between in-spec and out-of-spec aggregates. These wavelet-based features are used as inputs to an artificial neural network, which is used to assign a predefined class to the aggregate sample. Verification tests show that this approach can potentially help a plant operator determine, in a fast and accurate manner, if the aggregates currently being produced are in-spec or out-of-spec.

KEYWORDS: aggregate, artificial neural networks, group texture, laser profiling, wavelet transforms

1. INTRODUCTION

The importance of using high quality stone aggregates is gaining increased recognition within the construction industry. To rapidly acquire the data needed to ensure that aggregate products have the desired properties, automated methods for characterizing construction aggregates have been developed. By implementing automated methods of measuring basic material properties in testing laboratories, at large construction sites, and so forth, construction material quality can be improved.

Digital image analysis (DIA) has been widely studied as a means of automating aggregate tests [1]. In DIA, an aggregate sample from the production stream is photographed with a camera; this image is then digitized for computer analysis. To extract size information on each particle in the digital image, algorithms for image segmentation and size measurement are used. That is, after the particles in the image are separated by the segmentation algorithm, all of the particles are measured, one by one, in a computationally intensive manner.

However, if the application is primarily concerned with variations in the product rather than complete sample characterization, a much faster approach is possible. For example, in aggregate production plants, product gradation can be monitored by tracking variations in the percent passing a selected sieve size [2]. By monitoring variances in this one measure, plant operators can know when the production process needs to be adjusted. The method of extracting simple variance information facilitates faster analysis because it does not require the complete characterization of each

[1] Post Doc. Researcher, Department of Civil Engineering, The University of Texas at Austin, Austin, TX 78712, USA, hyoungkwan@mail.utexas.edu.
[2] Professor, Department of Civil Engineering, The University of Texas at Austin, Austin, TX 78712, USA, haas@mail.utexas.edu.
[3] Assistant Professor, Department of Civil Engineering, The University of Texas at Austin, Austin, TX 78712, USA, arauch@mail.utexas.edu.

particle following segmentation. In addition, this approach can potentially enable the plant operator to assess the properties of aggregates on a conveyor belt without acquiring discrete samples from the belt.

This paper proposes a neural network based quality control method for aggregate production. The Laser-based Aggregate Scanning System (LASS) [3], developed at the University of Texas at Austin, is used to acquire accurate three-dimensional (3D) data on stone particles. To generate meaningful features for input to the neural network, two-dimensional (2D) wavelet transforms are suggested for processing the 3D data. Aided by the multi-resolution feature of the wavelet transform, the neural network is expected to provide the necessary information for real-time quality control during aggregate production.

This paper begins with a literature review covering group texture, 2D wavelet transforms, and artificial neural networks. Then, an aggregate classification system is proposed for monitoring variations in an aggregate product stream. This system is focused on detecting variations in particle size distribution (gradation). Finally, experimental results and conclusions are presented.

2. LITERATURE REVIEW

2.1 Group Texture and Wavelet Transforms

In the machine vision field, texture is defined as "something consisting of mutually related elements" [4]. Namely, texture can mean a combination of texture elements and the relation between each element. In an attempt to identify the most suitable method for objectively quantifying the properties of an aggregate sample, machine-vision-based texture quantification (or classification) methods were investigated. These methods included the use of statistical moments, co-occurrence matrix, edge based method, Law's energy, surface based method, fractal geometry, mathematical morphology, and Fourier transform.

Wavelet analysis, where edges on various scales are detected and processed, is a method that belongs to the edge based texture quantification methods. A wavelet analysis decomposes a signal into a group of linear combinations, with each combination having different resolutions. This transform is conducted using the finite length of a basis function called a "mother wavelet". The mother wavelet is compared with the signal to be analyzed by changing its length (dilation) and location (translation) in order to find where and how much each dilated and translated version of the mother wavelet coincides with the signal. The dilation and translation mechanism of the mother wavelet enables not only production of localized information in the space and frequency domains, but also effective representation of the data signal.

A comprehensive explanation of 2D wavelet transforms can be found in [5,6,7]. With a 2D wavelet transform, a digital grayscale image can be represented as:

$$
\begin{aligned}
f(x,y) = & \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} c_0(j,k)\varphi(x-j)\varphi(y-k) \\
& + \sum_{i=0}^{\infty}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} d_{i,0}(j,k)\varphi(2^{-i}x-j)\psi(2^{-i}y-k) \\
& + \sum_{i=0}^{\infty}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} d_{i,1}(j,k)\psi(2^{-i}x-j)\varphi(2^{-i}y-k) \\
& + \sum_{i=0}^{\infty}\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} d_{i,2}(j,k)\psi(2^{-i}x-j)\psi(2^{-i}y-k)
\end{aligned} \tag{1}
$$

where $f(x,y)$ is the grayscale image, φ is the scaling function of the 1D wavelet transform, ψ is the wavelet of the 1D wavelet transform, j and k represent a location in the wavelet domain, i represents a decomposition level, and c_0 and $d_{i,l}$ (l = 0, 1, 2) are coefficients for a scaling function and wavelets, respectively.

2.2 Artificial Neural Networks

Artificial Neural Networks (ANNs) are pattern recognition systems that imitate biological nervous systems. ANNs can be used either as classifiers, to allocate a predefined category to the data representing a given case, or as estimators for predicting a certain value based on the given environment. A typical ANN consists of three different layers: the input layer, hidden layer, and output layer. While there is only one input layer and one output layer, the number of hidden layers used usually depends on the degree of complexity in the pattern recognition problem. Each layer has one or more processing elements called neurons (or nodes), which are typically connected with those of the next layer. These neurons take input signals, process them, and produce output signals. These signals are weighted and transferred using the connections between neurons.

To operate properly, ANNs must be trained with many examples. This study uses a backpropagation training algorithm, one of the simplest and most general methods for training multilayer neural networks [8]. In the backpropagation method, the network propagates the errors, determined by the differences between the actual and desired output values, backward (from the output layer to the input layer) while adjusting the connection weights between neurons. A more comprehensive treatment of ANNs can be found in [8].

3. PROPOSED METHOD

3.1 Laser-based Aggregate Scanning System

The "Laser-based Aggregate Scanning System" (LASS) was developed to acquire 3D aggregate particle surface data. The LASS consists of a laser line scanner, a horizontal gantry system, and a personal computer (Fig. 1). The laser scanner, which is mounted on the gantry system, passes over an aggregate sample, scanning it with a vertical laser plane. The laser line scanner can move approximately 1.5 m along the Y axis while performing 25 scans per second, with a scan width (X axis) of 120 mm and a scan height (Z axis) of 220 mm.

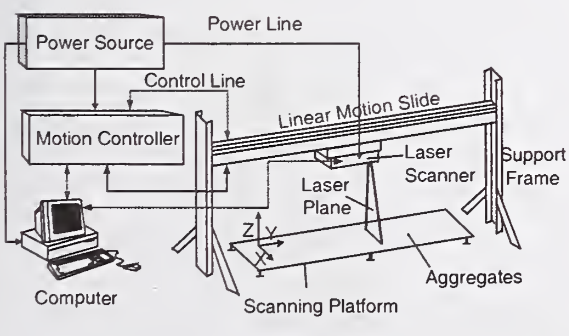


Figure 1. The Laser-based Aggregate Scanning System (LASS).

The resolution of the LASS data is as good as 0.3 mm, 0.1 mm, and 0.5 mm in X, Y, and Z directions, respectively. A comprehensive description of the LASS can be found in [3].

3.2 Artificial Intelligence Based Aggregate Classification System

Texture can be defined as a combination of texture elements and the relations between each element. Aggregate particles can correspond to texture elements with certain special relationships with each other. If a group of construction aggregates is scanned into an image, this image can be considered as a texture. One method to quantify texture uses edge information in the image. For example, the number of edge pixels in a certain area can be used for texture description.

Texture descriptions are highly scale dependent [4]. For instance, edges detected with high resolution would be ignored if low resolution was used. However, wavelet analyses can be used to advantage in overcoming this problem. 2D wavelet analysis provides vertical, horizontal, and diagonal edge information on various scales. With this information, it is possible to quantify the texture of an aggregate image effectively and objectively. Then, by comparing this quantified information between in-spec and out-of-spec aggregate images, an aggregate group with an out-of-spec gradation can be detected as unacceptable.

A flow chart for the proposed aggregate classification system is shown in Fig. 2. Aggregate samples are first scanned by the LASS to obtain 3D laser images. The height value of each data point is represented by a grayscale value ranging from 0 to 255. Then, 2D wavelet transforms are applied to the images so that the following features can be obtained:

$$\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{l=0}^{2}\left|d_{i,l}(j,k)\right| \tag{2}$$

Basically, the features are energies (summation of absolute values of all the elements) of the decomposition level *i*. Since particles are randomly spread and scanned, no distinction is necessary between horizontal, vertical, and diagonal edges in the wavelet transformed image. This is why d_0, d_1, and d_2 can be added together. In other words, all the edge information on a resolution level is summed to obtain one feature value, which is then put into a classifier to determine the appropriate categories for the aggregate sample. In this approach, the number of decomposition levels in the 2D wavelet transforms applied to the image is naturally the maximum number of features that can be used in the proposed classification system.

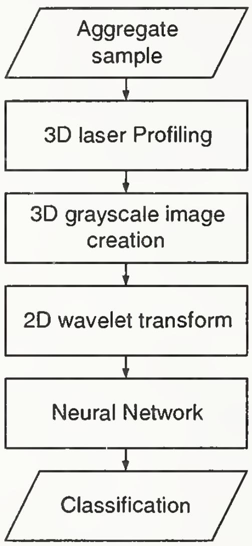


Figure 2. Artificial intelligence based aggregate classification system.

In this study, an artificial neural network (ANN) is used as a classifier to determine whether or not an aggregate sample is out-of-spec. Since this research is focused on detecting only variations in particle size distribution, the following three groups (categories) are defined: Norm, Large, and Small. Group Norm is composed of 100 % of the same size of aggregates (which would pass a certain mesh size and be retained on a certain smaller mesh size). Group Large has a certain percentage of larger particles and Group Small has a certain percentage of smaller particles. Thus, this system can classify an aggregate sample into three categories: in-spec, out-of-spec with larger particles, and out-of-spec with smaller particles.

In a field application, these classifications could be used to adjust the aggregate production process. If the plant operator finds that the aggregates currently being produced are classified as out-of-spec with larger particles, the crusher settings could be tightened to produce fewer oversized particles. If the categorization indicates out-of-spec with smaller particles, the crusher's settings could be opened to produce fewer small particles. Depending on the specific needs of the aggregate producing plant, more than three categories could also be defined.

Fig. 3 shows the neural network model for the aggregate classification system. It is composed of an input layer with two neurons, a hidden layer with five neurons, and an output layer with three neurons. The number of input features naturally determines the number of input neurons, while the number of output neurons is determined by the number of categories used to classify the aggregate samples. A sigmoid nonlinear function and backpropagation with a momentum learning method were adopted for training the neural network model.

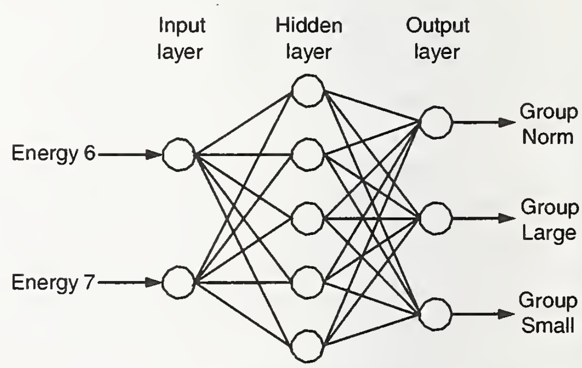


Figure 3. Neural network model for the aggregate classification system.

Sixth and seventh energy levels, which correspond to relatively low frequencies in the wavelet domain, are fed into the neural network model. These two features were selected because preliminary experiments indicated that those energy levels are most apt to differentiate between aggregate groups with different gradations. This preliminary examination of the energy features saves a significant amount of computing effort by reducing the complexity of the neural network. It is also worth noting that the network has three output neurons matching the three categories defined as Norm, Large, and Small, whereas the number of neurons for the hidden layer was determined from trial and error.

The aggregate classification system was implemented using the C++ programming language, LabView (a graphical programming language), the IMAQ Vision image processing tool, the Wavelet and Filter Bank Design Toolkit, and the DataEngine (an off-the-shelf Neural Network subroutine). LabView, IMAQ Vision, and the Wavelet and Filter Bank Design Toolkit are all products of National Instruments (Austin, Texas), while the DataEngine is a product of MIT GmbH (Germany).

4. EXPERIMENTS

To check the validity of the group texture and artificial intelligence based aggregate classification method, the proposed system was used to classify three aggregate samples described in Table 1. Norm particles, Large particles, and Small particles are defined as particles that fall within the size ranges of 1/2" to 3/4" (12.7 mm ~ 19.0 mm), 1" to 1-1/4" (25.0 mm ~ 31.5 mm), and No. 4 to 3/8" (4.75 mm ~ 9.5 mm), respectively. Then, Group Norm consists of 100 % of Norm particles, Group Large has 50 % of Large particles and 50 % of Norm particles, and Group Small has 50 % of Small particles and 50 % of Norm particles. These aggregate samples were randomly spread on the scanning platform of the LASS such that there are no overlapping particles. They were then scanned and converted into digital images. Fifty-six images were created for each group, resulting in a total of 168 images. Each image is 566 by 180 pixels and covers a rectangular area of 120 mm by 50 mm. Eighty-four images (half the total number of images) were used to train the neural network model described in Fig. 3, while the other 84 images were used to test the classification system.

To obtain the energy values that are representative of each aggregate sample, a running (moving) average value of every five images was used instead of separate energy values for each image, as follows:

$$RA_i = \frac{1}{5}\sum_{j=i}^{i+4} I_j \qquad (3)$$

where *RA* is a running average and *I* is an energy feature of one image. In other words, for every five images in the same aggregate group, the energy values were averaged to produce more stable and representative features. Note that this running average approach reduced the total number of training sets (or test sets) from 84 to 72.

Table 1. Description of aggregate test samples.

Group	*Size fraction*	*%*	*kg*
Group Norm	No. 4 ~ 3/8" (4.75 mm ~ 9.5 mm)	0	0
	1/2" ~ ¾" (12.7 mm ~ 19.0 mm)	100	5
	1" ~ 1-1/4" (25.0 mm ~ 31.5 mm)	0	0
Group Large	No. 4 ~ 3/8" (4.75 mm ~ 9.5 mm)	0	0
	1/2" ~ ¾" (12.7 mm ~ 19.0 mm)	50	2.5
	1" ~ 1-1/4" (25.0 mm ~ 31.5 mm)	50	2.5
Group Small	No. 4 ~ 3/8" (4.75 mm ~ 9.5 mm)	50	2.5
	1/2" ~ ¾" (12.7 mm ~ 19.0 mm)	50	2.5
	1" ~ 1-1/4" (25.0 mm ~ 31.5 mm)	0	0

Table 2 shows the classification results. With only one incorrect classification in identifying Group Large, the 99 % classification accuracy demonstrates that the group texture approach, in conjunction with artificial intelligence classifiers, is a promising method to detect variations in an aggregate production stream.

Table 2. Classification results.

Group	*Accuracy (Number)*	*Accuracy (%)*
Group Norm	24 / 24	100
Group Large	23 / 24	96
Group Small	24 / 24	100
Total	71 / 72	99

5. CONCLUSION

This paper explored the possibility of using group texture of aggregate images in conjunction with an artificial neural network to quantify gradation properties. First, an aggregate sample was scanned by the Laser-based Aggregate Scanning System, and converted into 3D images. Then, 2D wavelet transforms were applied to those images to extract wavelet coefficients and calculate energies at different scales. Finally, these energies were used as inputs to an artificial neural network that assigns a predefined class to the aggregate sample. Verification tests show that this approach can potentially classify aggregates in a fast and accurate manner.

Further work is needed to develop and verify the proposed artificial intelligence based approach. First, while reducing gradation variations in the training samples, different network architectures can be constructed and evaluated to optimize the neural network model. This requires testing with different numbers of neurons, different numbers of hidden layers, different transfer functions, and different learning methods. Second, efforts are needed to develop good features that can represent the aggregate properties well. A system using the standard deviation or other statistics of the wavelet coefficients at a certain decomposition level might be successful in grouping similar aggregate samples. Third, different classifiers, such as the K-Nearest-Neighbor method, Linear discriminant function, Fuzzy logics, etc. [8], could be investigated. These relatively simple methods are sometimes more effective than complicated neural networks.

6. ACKNOWLEDGMENT

Funding for this study was provided by the International Center for Aggregates Research (ICAR).

7. REFERENCES

1. Browne, C., Rauch, A. F., Haas, C. T., and Kim, H. (2001). "Comparison tests of automated equipment for analyzing aggregate gradation." *Proc., 9th Annual Symposium,* ICAR, Austin, Texas.
2. National Stone Association. (1993), *The Aggregate Handbook*, Mercury Press, Rockville, MD.
3. Kim, H., Haas, C. T., Rauch, A. F., and Browne, C. (2001). "A prototype laser scanner for characterizing size and shape parameters in aggregates." *Proc., 9th Annual Symposium,* ICAR, Austin, Texas.
4. Sonka, M., Hlavac, V., and Boyle, R. (1999). *Image Processing, Analysis, and Machine Vision*, PWS Publishing, Pacific Grove, CA.
5. Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS series, SIAM, Philadelphia, PA.
6. Mallat, S. (1999). *A Wavelet Tour of Signal Processing, Second Edition*, Academic Press, San Diego, CA.
7. Rao, R. M., and Bopardikar, A. S. (1998). *Wavelet Transforms*, Addison Wesley Longman, Inc., Reading, MA.
8. Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*, John Wiley & Sons, Inc., New York, NY.

A MEMS Transducer for Ultrasonic Flaw Detection

by

Akash Jain, David W. Greve, and Irving J. Oppenheim[1]

ABSTRACT

Metal structures can fail because of fatigue crack propagation or because of section loss from corrosion. Regular inspection is required to intercept such failures, and *in situ* sensors would be a superior technology for that purpose. We have designed and fabricated arrays of MEMS capacitive diaphragm transducers and we report on their performance as pulse-echo detectors in direct contact with solids. Our chip is approximately 1-cm square and features nine detectors in a linear array, each detector containing 180 hexagonal diaphragms. Performance of the detector array was studied by bonding the chip to test specimens and applying an ultrasonic pulse using a commercial ultrasonic transducer. One experiment recreates an on-axis excitation in which the pulse arrives uniformly at all detectors, and another experiment recreates an off-axis excitation in which the pulse arrival is delayed from one detector to the next along the length of the array, permitting accurate localization of the source using phased array signal processing. The results establish that MEMS transducers can function successfully as phased array detectors of ultrasonic signals in solids.

KEYWORDS

Diaphragm; flaw detection; MEMS; phased array; ultrasonics

1. INTRODUCTION

Steel is used in buildings, bridges, pressure piping, and industrial construction, but the safe performance of any steel structure is threatened by section loss from corrosion or wear, by crack propagation from fatigue or cyclic loading, by weld failure from overload or seismic loading, or by other discontinuities. Such flaws can develop with time, and the continued service of major structures often requires confirmation that such flaws have not developed. Ultrasonic flaw detection [1] is a versatile technology for nondestructive evaluation, but it must typically be performed by skilled personnel. The principles of pulse-echo flaw detection are depicted in a through-thickness geometry in Figures 1 and 2. Figure 1 depicts an ultrasonic pulse transmitted into the material. A typical transducer frequency is 5 MHz, corresponding to a 1.2-mm wavelength in steel, which is sufficiently short to resolve flaws at that same scale. The typical transducer is a piezoelectric ceramic, most often PZT (lead-zirconium-titanate), with a diameter much greater than the wavelength. The ultrasonic pulse will reflect from the first boundary it encounters, which in an unflawed specimen is the back surface of the steel plate. The time at which the echo returns to the front surface reveals the total travel distance, equal to twice the thickness. Figure 2 records a measurement using a mm-scale PZT sample affixed to brass (velocity of sound c = 4400 m/s) with a thickness of 9.8 mm, showing successive echo returns. The time from the pulse to the return of the echo, and the time between successive echoes, is under 5 μs, which correctly approximates the thickness. Ultrasonic inspection can be used in this manner to measure thickness, which would reveal any section loss, or to reveal reflections that arrive prematurely, which would signal the presence of a flaw.

[1] Department of Civil and Environmental Engineering (IJO), Department of Electrical and Computer Engineering (DWG, AJ), Carnegie Mellon University, Pittsburgh, PA 15213; Phone 412-268-2950 (IJO); Email ijo@cmu.edu.

Although ultrasonic flaw detection is quite versatile, there are two limitations that could be eliminated by the development of resident sensors. In current practice the inspection is performed manually, and is therefore subject to interpretation. Moreover, the process is most often memoryless, making no use of the earlier signal history. We envision building a resident ultrasonic flaw detection system to be mounted at critical locations on metal structures, which would retain a signal history to allow signature analysis in the detection of developing flaws. We intend that the device be polled remotely using RF communications. This paper describes the design and initial testing of a MEMS capacitive (diaphragm-type) transducer array to function as the receiver in the flaw detection system. In order to scan a volume of material from a fixed position it is necessary for the transducer to function as a phased array, and experiments to demonstrate signal detection and phased array processing were a main purpose of this study.

2. PREVIOUS WORK

Ultrasonic pulse-echo detection is used in many applications including range/motion sensing, embedded object detection, surface characterization, and medical ultrasound imaging. There is a considerable history of research into MEMS transducers for fluid-coupled and air-coupled applications. Our approach to designing microscale ultrasonic diaphragms was based on the important work of Khuri-Yakub at Stanford University [2,3,4]. One paper [2] outlines the mechanical and electrical analysis of capacitive diaphragm transducers and presents experimental results for air-coupled and fluid-coupled transmission through aluminum, showing that practical applications (including flaw detection) are feasible. Another paper [3] records in detail the fabrication steps needed to produce capacitive ultrasonic transducers suitable for immersion applications and the characterization, both experimental and analytical, of their performance. Another reference [4] discusses one-dimensional transducer arrays and presents initial imaging results, in which solids immersed within fluids are detected. Other investigators of MEMS ultrasonics include Schindel [5] with numerous contributions to immersion applications, and Eccardt [6], at Siemens, with the demonstration of surface micromachined transducers in a modified CMOS process. The present authors [7] have recently published an earlier version of the experimental results described herein.

3. DEVICE DESIGN

In a MEMS capacitive transducer, a DC bias voltage is maintained across the plates of the capacitor and diaphragm deflection then produces a change in capacitance that can be detected electrically. The sensitivity of a single diaphragm increases linearly with its area and with the bias voltage, and inversely with the cube of the gap dimension. Moreover, the sensitivity of a detector composed of diaphragms in parallel increases with the number of diaphragms, and therefore a favorable utilization of area is preferred in order to obtain maximum signal strength. Accordingly, a hexagonal geometry was chosen for the individual diaphragm unit and the transducer was fabricated by the MUMPS surface micromachining process. The diaphragm is constructed in the polysilicon-1 structural layer with a thickness of 2 μm and is a regular hexagon with leg length equal to 49 μm, chosen to yield a resonant frequency near 4 MHz. A target capacitance of a few pf was chosen, but the predicted capacitance for a single diaphragm was only 0.016 pf; therefore the basic detector was fabricated as a group of 180 diaphragm units in parallel. Figure 3 is the layout drawing for a typical detector, with approximate dimensions of 0.9x2 mm.

The overall device layout is shown in Figure 4. The chip is approximately 1-cm square and contains 23 detectors. The primary detector array is the set of nine in the right-hand column, spanning a 1-cm baseline for phased array implementation. The nine detectors in the middle column are an alternate design attempting to use the substrate, rather than a deposited electrode surface, as the stationary plate of the capacitor. The three detectors at the top of the left-hand column constitute variations on the diaphragm design, using closer-spaced etch release holes, to perform experiments on squeeze film damping. The two largest detectors in the left-hand column are alternate diaphragm designs constructed with two polysilicon layers, for a thickness of 4 μm, and a correspondingly larger leg dimension of 69 μm.

4. EXPERIMENTAL RESULTS

To our knowledge, our tests were the first to attempt signal detection by MEMS transducers in direct contact with solids. The experiments were performed with chips bonded to plexiglass specimens using Gelest Zipcone CG silicone adhesive. Commercial ultrasonic transducers, with nominal diameters of 15 mm and rated operating frequencies of 3.5 MHz and 5 MHz, were the signal sources. Figures 5a and 5b depict the specimen geometries; the MEMS chip appears on-edge as a small rectangle, and the dimension records the closest distance between the signal source and the nearest detector. In the test depicted in Figure 5a the baseline of nine detectors appears as a single point, the point closest to the transducer. Because the transducer is approximately 15 mm wide and the detector baseline is less than 10 mm long, the signal is expected to arrive simultaneously at each detector; the test is termed the on-axis geometry. In the test depicted in Figure 5b the baseline of nine detectors appears as the heavy line. The dimension shown (0.72-inch, or 18 mm) is the distance between the signal source and the nearest single detector. Therefore, the signal will reach the nine detectors along the baseline at an extreme raking angle (65° from the normal) and with considerable delay in arrival time across the baseline; the test is termed the off-axis geometry. The main purpose of these tests was to obtain the distance and angle between the transducer and the source in Figure 5b, using phased array signal processing.

Figure 6a shows experimental results for a pulse in the on-axis geometry illuminating the array of nine detectors from a distance of approximately 0.53-inch, or 13 mm. The signal received at each detector is displayed on the plot at the appropriate relative spatial position of each detector, and we note the following:

- Each signal shows a pulse near 1 μs because of stray electrical coupling, followed by the signal arrival approximately 4.5 μs later, corresponding roughly to the specimen thickness along that travel path.
- As predicted, the arrival time is uniform at all detectors.
- The signals at each detector are relatively uniform in appearance and comparable in amplitude.

Figure 6b shows experimental results for a pulse in the off-axis geometry raking across the array of detectors, and we note the following:

- Only seven detectors are shown, because two detectors became non-operative during the course of the tests.
- The signal arrives first at the closest detector, with successive delay in its arrival at each subsequent detector.
- The arrival times are consistent with the distance between the pulse source and the array.
- The delay permits localization of the source, determining the distance and angle to that source, using the principles of radar imaging.
- A simple geometric localization can be envisioned directly on Figure 6b. If a vertical line is drawn through the start of the pulses, and another straight line is drawn through the start of the received signals, those lines will intersect at a position that can be scaled (either from the inter-detector spacing or from the whole baseline dimension) to obtain the location of the pulse origin to the "left" of the array as it appears in Figure 5b.

A simple signal processing approach was used. Because the detectors are equally spaced, the delay between them will be roughly constant. If each signal is shifted successively by some delay, and then all signals are added together, the sum should be maximum when the correct delay is used. Equivalently, "guessing" the distance and the angle to the source constitutes a "guess" at a delay, with which the signals can be summed, and when the best estimates of distance and angle are used the sum should be a maximum. Figure 7 depicts the results of that process, using arbitrary units, and the isolated peak represents the best estimate of distance and angle to the source; the axis projecting into the foreground represents the distance and the other axis right represents the angle to the source. (The peaks along the distance axis represent the stray-coupled pulses, and should be ignored.)

5. CONCLUSIONS

Experimental results in Figures 6 and 7 show that MEMS capacitive (diaphragm-type) transducers can successfully detect ultrasonic pulses when in contact with a solid. The phased array implementation shows that the transducer can successfully localize a source in an off-axis geometry. This first-generation device was designed to test the feasibility of phased array detection, to evaluate design alternatives, and to conduct related experiments in diaphragm behavior. The detectors fabricated in this first device are sufficiently sensitive to detect pulses from a commercial PZT transducer. More recent results (not shown) demonstrate that the detectors are sufficiently sensitive to detect pulses from mm-scale PZT sources if geometric spreading from the signal source is kept small. However, demonstration of flaw detection in practical geometries will require greater sensitivity in order to detect signals from small sources (creating a spherical wave) after considerable geometry spreading. Currently the sensitivity is limited by the capacitor gap and the detector area, and detection limits are severely constrained by parasitic capacitances. A second-generation device is presently being fabricated with a number of design improvements to these conditions, and is expected to improve performance by an order of magnitude. Additional improvements in effective sensitivity, by orders of magnitude, can be achieved when the mechanical transducer and the electronic circuits are fabricated as an integrated system on a single chip.

6. ACKNOWLEDGEMENTS

This work has been funded by the Commonwealth of Pennsylvania through the Pennsylvania Infrastructure Technology Alliance program, administered at Carnegie Mellon by the Institute for Complex Engineering Systems, and by gifts from Krautkramer Inc., and the authors gratefully acknowledge that support.

7. REFERENCES

1. Krautkramer, J. and Krautkramer. H., *Ultrasonic Testing of Materials*, 4^{th} edition, Springer Verlag, Berlin, 1989.

2. Ladabaum, I., Jin, X., Soh, H., Atalar, A., and Khuri-Yakub, B., “Surface micromachined capacitive ultrasonic transducers,” *IEEE Trans. Ultrasonics,* Vol. 45, 678-690, 1998.

3. Jin, X., Ladabaum, Degertkin, F., Calmes, S., and Khuri-Yakub, B., “Fabrication and characterization of surface micromachined capacitive ultrasonic immersion transducers,” *IEEE Jnl. Of Microelectricmechanical Systems*, Vol. 8, 100-114, 1999.

4. Oralkan, O., Jin, X., Kaviani, K., Ergun, A., Degertkin, F., Karaman, M., and Khuri-Yakub, B., “Initial pulse-echo imaging results with one-dimensional capacitive micromachined ultrasonic transducer arrays,” *IEEE Ultrasonics Symposium*, 959-962, 2000.

5. Schindel, D., “Air-coupled generation and detection of ultrasonic bulk-waves in metals using micromachined capacitive transducers,” *Ultrasonics*, Vol. 35, 179-181, 1995.

6. Eccardt, P., Niederer, K., Scheiter ,T., and Hierold, C., “Surface micromachined ultrasound transducers in CMOS technology,” *IEEE Ultrasonics Symposium*, 959-962, 1996.

7. Jain, A., Greve, D. W., and Oppenheim, I. J., "A MEMS Ultrasonic Transducer for Monitoring of Steel Structures," *SPIE Smart Systems for Bridges, Structures, and Highways*, Paper 4696-31, San Diego, March 2002.

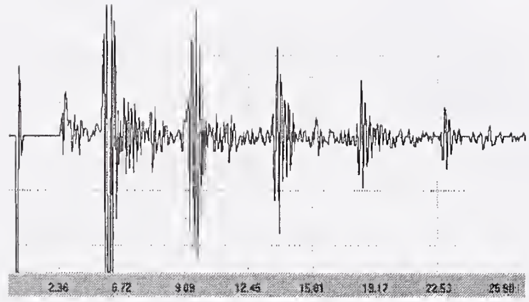


Figure 1. Pulse-echo flaw detection, ref [1]   Figure 2. Results using mm-scale PZT specimen

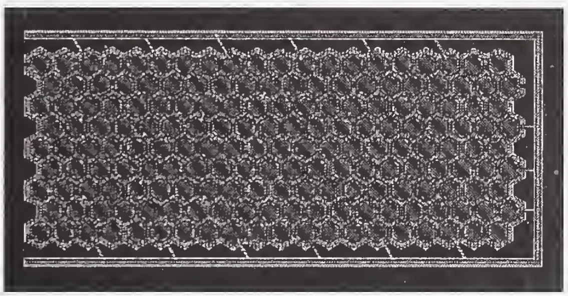


Figure 3. Typical detector, approximately 0.9x2mm, containing 180 diaphragms

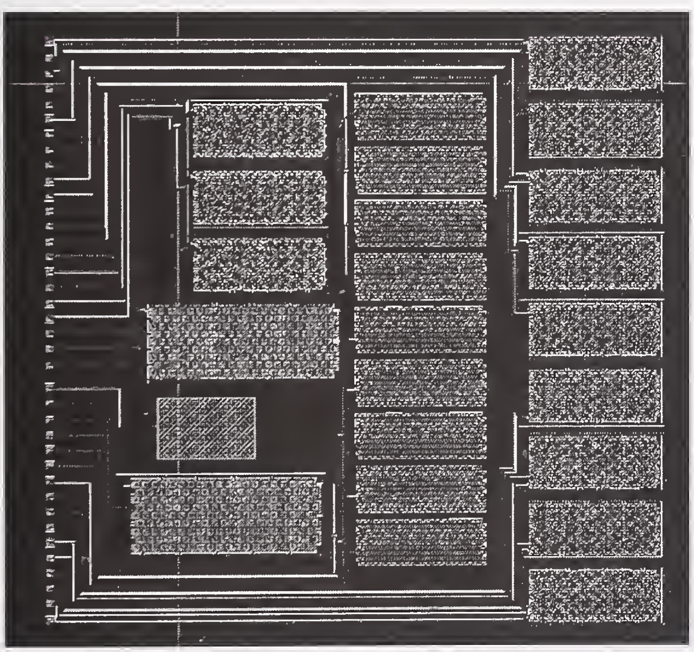


Figure 4. Layout drawing of MEMS chip, array of nine detectors in right-hand column

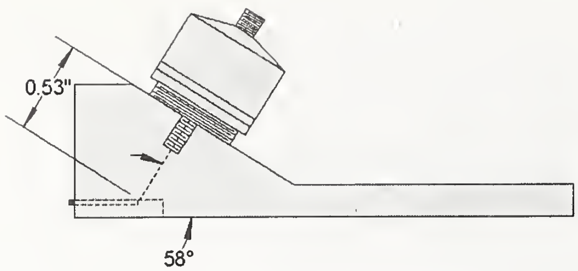


Figure 5a. Test specimen, on-axis geometry

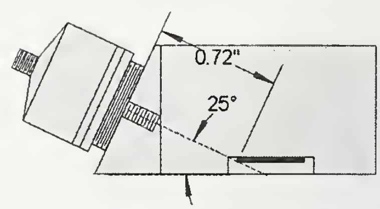


Figure 5b. Test specimen, off-axis geometry

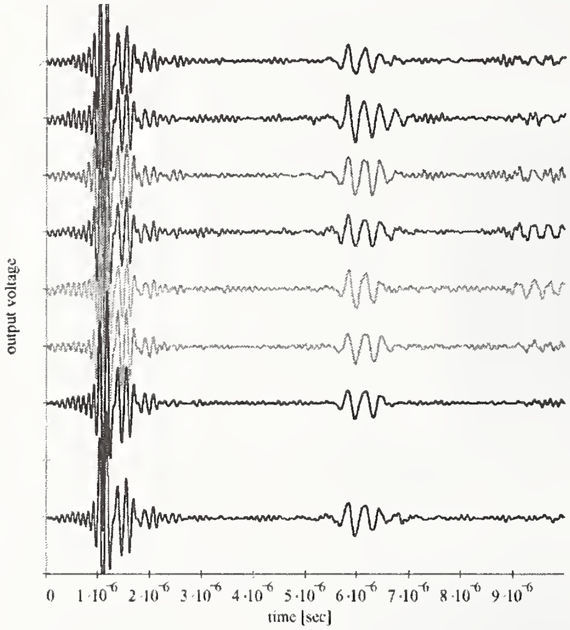


Figure 6a. Experimental results, on-axis geometry

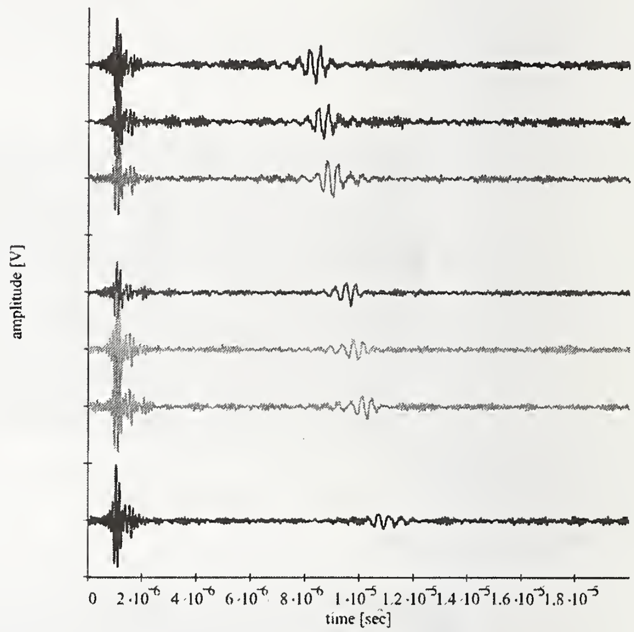


Figure 6b. Experimental results, off-axis geometry

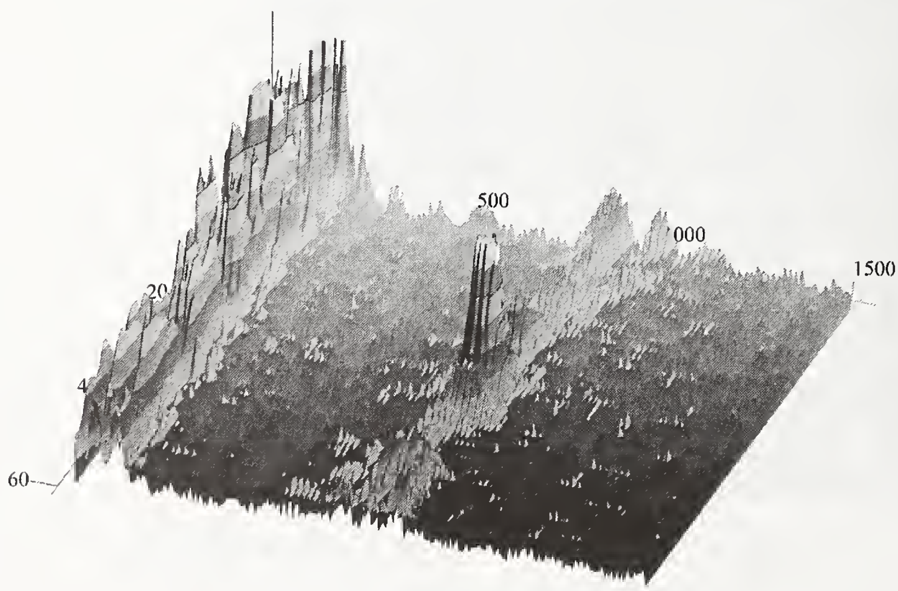


Figure 7. Signal processing results, distance and incidence angle to source given by the peak

Utilizing Radio Frequency Identification on Precast Concrete Components – Supplier's Perspective

Burcu Akinci,[1] Mark Patton[2], Esin Ergen[3]

ABSTRACT: Precast concrete material suppliers are responsible for the components that they manufacture from casting until up to 25 years after installation. To effectively manage the production and storage at a the production facility, to streamline the construction process and to quickly repair a component should there be a problem, information about components must be readily available and updated throughout the 25 year life cycle. Currently, suppliers use barcodes to track precast components and paper-based documents to store information about them. These approaches are time-consuming and error prone. Suppliers also face the problem of not finding the right information in a timely manner. This paper discusses the utilization of radio frequency identification technology (RFID) for tracking precast concrete components and their historical information from fabrication to post-construction. RFID is an automatic identification technology that uses memory chips attached/embedded to objects to transmit data about them. Our discussion revolves around one use case that we developed describing the current component tracking approach and the proposed approach utilization of RFID technology from a large-scale precast manufacturer/erector's perspective. We conclude with an assessment of benefits and limitations of the proposed utilization of RFID system for tracking precast concrete materials.

KEYWORDS: RFID; precast concrete; construction; material management

1. INTRODUCTION

A precast concrete manufacturer is liable for a component from the time that they cast the component until 25 years after installation. Precast components used in a facility are expected to maintain their original condition throughout their life cycles. Any problem that might have occurred during the manufacturing and erection of a component might not surface until late in service. Once a problem is detected during service, it is important to know the full history of a component to address the issue effectively.

At the manufacturing phase, the material information about precast components should be documented during casting. After the components are cured the information on where they are being stored should be documented to minimize the time-loss in locating components in a large storage area. At the construction phase, deliveries should be inspected to ensure that the components have the desired quality. Any problems during installation should be documented as part of a component's history.

Currently, critical information about fabrication and installation processes is not stored at all or cannot be accessed easily after the completion of a project, since they are mostly paper-based and stored at different locations. As a result, when a problem occurs with a component, it is hard to access the relevant historical data important to identify the cause of a problem.

Part of material tracking problem, especially at the site or a large storage area is being addressed by barcodes (Bernold 1990). Even though barcodes provide an easier way to track components, it has been noted that barcodes are being damaged in the harsh construction environment. In addition, in some cases, it is difficult to access the tag

[1] Assist. Prof., Civil and Env. Eng. Dept., Carnegie Mellon University, Pittsburgh, PA, 15213, bakinci@andrew.cmu.edu

[2] Research Prof., Civil and Env. Eng. Dept., Carnegie Mellon University, Pittsburgh, PA, 15213, mp3c@andrew.cmu.edu

[3] Graduate student, Civil and Env. Eng. Dept., Carnegie Mellon University, Pittsburgh, PA, 15213, eergen@andrew.cmu.edu

for scanning (especially in storage areas) since one needs to be next to a barcode to be able to read the information (Furlani 2000). Finally, barcodes do not have any memory and hence it would be impossible to store the history of a component on a barcode.

Radio Frequency Identification (RFID) is another identification technology that uses memory chips attached or embedded to objects to transmit data about them. In this paper, we discuss the utilization of RFID technology to locate and track the precast pieces and to store information about them from their fabrication to installation. Our discussion revolves around a use case developed based on data obtained from a large-scale precast manufacturer.

2. BACKGROUND ON RFID

RFID is an automatic identification technology used to identify, track, and detect various objects. RFID systems are composed of two components: a tag, and a reader. RFID tag is an electronic label that stores data and is attached to objects. Readers, which send RF (Radio Frequency) signals for communication, are used to read data from these tags. A reader is composed of an antenna, a transmitter/receiver and decoder.

Current RFID technologies use three frequency ranges: low (100-500 kHz), intermediate (10-15 Mhz), and high (850-950 Mhz / 2.4-5.8 Ghz). Table 1 lists some of the unique characteristics of each frequency range. Low and intermediate frequencies are commonly used for inventory control, while high frequencies are used for railroad car monitoring and toll collection systems (AIM 2002).

RFID tags can be classified as either active or passive based on the power source. An active tag has an internal battery for power. A passive tag utilizes the energy generated by a reader/antenna. Active tags have a greater read/write range (up to 30 m.). However, they are larger in size, more expensive, and have a limited life span (5-10 years). Passive tags are cheaper, smaller, lighter, and have unlimited life span. However, they require a more powerful reader and have shorter read ranges.

Table 1. Characteristics of frequency bands

Frequency	Characteristics
Low 100-500Khz	Short to medium reading range Inexpensive Low reading speed
Intermediate 10-15 MHz	Short to medium reading range Potentially inexpensive Medium reading speed
High 850-950 MHz 2.4-5.8 GHz	Long reading range High reading speed Line of sight required Expensive

Tags also can be read only (RO), read / write (R/W) or write once / read many (WORM) (Paolo 2002). RO tags are pre-programmed with unique information, and this data cannot be modified afterwards. In R/W tags, additional data can be written by overriding or extending the existing data stored in a tag. In WORM tags, information in a tag can be changed only once. After the data is read from any type of transponder, it can be sent to a host computer, or stored on a reader to be later uploaded to a computer (AIM 2002).

Tag cost, size and memory capabilities vary. Tag information can be limited to a unique code, or include detailed information, e.g., manufacturer, storage conditions, etc. A passive tag can store up to 2k bits of user information (TI 2001) while an active tag can have a memory up to 1MB (AIM 2002). Price of a tag increases with the amount of information that can be stored on a tag. Current prices are between 50 cents to $50 depending on the memory, type (active/passive, RO, R/W or WORM) and operating frequency. However, prices are decreasing as the technology develops. For example, in 1995 prices for various tags were between $10-$100 (Jaselskis 1995).

RFID tags are technically more advantageous than barcodes in material tracking and documenting the history of a component. RFID tags have capability of larger data storage. In addition, they can operate in harsh environments and do not require line of sight like barcodes. Besides, it is much faster to collect information about a batch of components using RFID since RFID technology allows collection of up to forty items per second (Marconi 2002). Data security is achieved in RFID tags with an optional password mode, which requires a password to enable read and write functions. Data in a tag can be locked by the user to prevent future modification (Atmel 2002).

3. PREVIOUS EXAMPLES OF RFID USAGE AT CONSTRUCTION SITES

Several applications of RFID in other sectors include tracking pallets, mail, and automation processes in manufacturing. Schell (2001) suggests that the utilization of RFID tags during the maintenance of a plant's pressure relief valves has reduced manual data entry by 70%.

Jaselskis (1995) proposed some potential applications, such as, cost coding for labor and equipment, and material control, which utilizes RFID technology during a construction project. He identified three major barriers for the utilization of RFID technology at construction sites: (1) lack of standardization among different RFID manufacturers impeding interchangeable usage of RFIDs from different brands; (2) hampering effect of metals caused by the reflection of RF energy from metal surfaces; (3) requirement for a battery management program for the active tags (Jaselskis 2000). Today, current technology has overcome the metal effect problem (Forster 2002), and various industry and standards organizations are working on RFID standards.

4. USE CASES

In this section, a potential application of RFID in precast components is discussed from a supplier/erector's perspective. The data presented is based on real data obtained from a large-scale precast manufacturer.

The large-scale precast manufacturer that we studied works on design-build arrangement, where an architect designs a facility, and then through a bidding process selects a manufacturer to design, engineer, fabricate and erect the structure. Precast components vary from structural beams to miscellaneous items such as columns, and curved sections. Their production is 17,200 pieces per year.

The manufacturer's storage area is approximately 60 acres, and the average number of precast elements stored at the production facility is 3,500 pieces at a given time. After casting, the components are sent to the storage area after one or two days maximum. Once at the storage area a precast component is either shipped within a few days or going to be stored there for months. The main reason for long storage periods is that the manufacturer prefers to fabricate similar pieces in a casting bed at the same time. Thus, some pieces can be produced months in advance since they are similar to the ones that will be delivered in a week. The locations of ones that are stored for longer durations are typically changed several times before their delivery date.

4.1 Current Material Tracking Process

Currently, barcodes are used to identify each piece, along with a piece mark that is written on one end with marker. Since thousands of different pieces are stored in the storage area and line of sight is required for barcodes, finding the necessary component becomes very difficult. Currently, about 10 workers work at the storage area, and it takes about 30 minutes to 1 hour to locate a component in that large area. In some cases, when the workers cannot locate some pieces at a storage area, those pieces have to be cast again, resulting in late deliveries. Currently, the manufacturer pays approximately $60,000 for annual penalties for late deliveries, which are mainly caused by delays associated with locating materials.

After a delivery is made, a field engineer is required to inspect each delivery to check if the right pieces are delivered. Then, the precast pieces are stacked in lay down areas for preferably 1-2 weeks (Pheng 2001).

In addition to tracking storage and delivery of precast pieces, precast manufacturers are also liable for any possible defects and failures of the precast pieces for up to 25 years. Different factors, e.g., cement, aggregates, the number of times a component is handled, the site conditions during construction, etc., can cause a defect on a precast component. When there is a defect or failure, they need to quickly identify these factors that the defective component is being exposed to so that they can fix the problem efficiently and prevent any similar failures in other pieces. Currently, much of the information related to the history of a component is not stored at all, and the ones, which are stored in documents, are not easily accessible.

In summary, precast suppliers/builders face two major problems associated with component tracking: (1) locating the precast pieces at their storage facilities and (2) accessing a component's history after construction. Current approaches for locating components and accessing a component's history are time-consuming, costly and not effective. RFID technology provides a way to address both of these problems. The next section describes the proposed system.

4.2 Proposed System

The proposed system utilizes RFID technology to locate each precast piece on a storage area without line of sight requirement, and ensure a reliable delivery schedule to a construction site. Additionally, information about manufacturing, inspection, and construction can be entered to a tag.

When an owner places an order, the designers in the company designs the components, and assigns an ID for each piece. During the fabrication of each piece, passive R/W tags with unique IDs are attached to the pieces. During that time all production related information is entered to the tag (Figure 1). After 1-2 days when the curing is complete, the pieces are checked by the inspector. The inspector's approval about the quality of precast components will also be stored on the tag before the components are sent to the storage area.

A grid of transponders ("tags") are embedded in the floor of a storage area. When a piece is placed in a storage area, first its tag is read to retrieve its ID and all the other information, and then the embedded transponders are read to match the transponder's code with the location in the storage area. This information in the reader is then sent to a supplier's material database. This would enable quick identification of the member at the storage area (Figure 2) and hence just-in time delivery of the pieces to a construction site. Whenever the location of a piece is changed, the same process is repeated, and any new handling information will be stored in the tag as part of the history of a component.

Once components that need to be delivered are identified, they are loaded onto a truck, which is also tagged with a W/R transponder. The IDs of the pieces are read and entered to truck's tag by a handheld reader. When the truck departs, a reader installed at the exit records the components that are shipped. The date, truck ID, component IDs are sent to the supplier's material database, and to the scheduler who works for the erector. Receiving this information, staging areas are determined on a site, and a crane is assigned for unloading.

Additionally, during transportation of materials, the delivery truck can be tracked by GPS; thus, both the manufacturer and the erector can be informed about the real-time location of the truck and the components.

At the entrance of a construction site, truck ID is read, components recorded as arrived in a database, which can be used for invoices and payments, and the manufacturer is notified that the delivery is completed. Another inspector who works for the erector company, checks the pieces, and enters any damage or defect that might have occurred during transportation. Finally, pieces are installed according to their Ids, and the current site conditions, and information about installation methodology are also stored in the tag as part of the history of components

As a result of this new process, delays due to locating materials at a manufacturing storage area and on a construction site are minimized and all the important information about the history of a component is stored on a tag. This information can be retrieved any time after the installation of a component should there be a problem. As a result, a manufacturer can efficiently identify the cause of the problem, effectively manage the problem, and take precautions for similar problems that can be observed in other components.

5. COST / BENEFIT ANALYSIS

The following cost/benefit analysis attempts to assess costs and benefits of locating materials on in a large storage area using RFID. The analysis is based on the quantitative data obtained from the manufacturer and on the current RFID costs. It is assumed that all precast components are tagged with a passive R/W tag, and workers are using three readers (Table 2).

Table 2: RFID equipment cost for one year

	number	Cost ($)
tag	17,200	172,000
reader	3	6,000
Total		178,000

As Table 2 shows, the cost of embedding tags and retrieving information from the tags is $178,000. To recover this cost, nominally 7,120 hours of worker time should be eliminated as seen in Table 3. We believe that using tags will save at least half of the actual time spent on locating the components. This corresponds to approximately 5 workers, which is 9,600 hours of worker time. This exceeds the 7,120 hours required to nominally recover the cost of implementing the RFID system and hence suggests that the savings will outweigh the cost of utilizing RFID tags

Table 3: Worker time need to be saved

Tag cost ($)	Hourly rate for a worker ($)	Worker time (h)
178,000	25	7,120

The cost benefit analysis based on the quantitative data shows that it is beneficial to use RFID for material tracking purposes. Also, there will be additional savings associated with just in time (JIT) delivery of components. Since currently the cost of delayed delivery is $60,000 per year, the JIT delivery of the components will enable additional savings up to that amount.

The proposed analysis, however, does not consider the potential additional benefits that will be received when the tags are also used for storing historical information of a component as described in Section 4.2. It is foreseeable that additional benefits will be incurred when same tags are used for storing historical information. Current memory capacity of a R/W passive tag, which is around 2k bits, might not be sufficient to store all information necessary. In the next phase of this research, we will be performing tests to assess how much of the necessary historical information. Regardless, the technology is developing rapidly, and memory capacity is likely to increase significantly in few years. With that, additional benefits associated with storing historical information will be achieved. This will create significant savings to the precast suppliers since currently majority of their consultants' time is spent in collecting historical data rather than analyzing it to repair a component.

In the future, when the memory capacity increases, the digital picture of any defect can also be stored in the tag.

6. CONCLUSION

In this paper, we propose a system, which utilizes RFID technology to locate precast components in a storage area at the manufacturing plant, tracks the delivery of the components, and stores all the necessary information about the components for the future use. In this system, passive RFID tags are suggested to be used for storing data, and for tracking precast pieces.

The system improves the efficiency of locating the components in storage areas in manufacturing plants. In addition, tracking the components during the delivery process is beneficial for just-in-time delivery. Finally, storing the necessary information in the tag helps the manufacturers to retrieve the history of a component quickly and identify the cause of any problem that might occur after construction.

Our future work includes the implementation of this system in a precast manufacturing plant. Through that, we will further identify the technological capabilities of current RFID tags in storing historical information and develop a more comprehensive cost/benefit analyses for the proposed use case.

7. REFERENCES

Association for automatic identification and data capture technologies (AIM) : http://www.aimglobal.org/technologies/rfid/what_is_rfid.htm, 06/27/2002.

Atmel Corp.: http://eu.atmel.com/atmel/acrobat/doc2003.pdf, 06/27/2002.

Bernold, L.E.: "Barcode-driven equipment and materials tracking for construction", J. of Comp. in Civil Eng., 1990, Vol. 4, No. 4, 381 –395.

Forster, I.J: "Resolving Problems in Using RFID Tags with Metal Objects", http://www.isit.com/feature.cfm?tech=AIDC&ArticleID=3829, 06/27/2002.

Furlani, K. M., Pfeffer L. E : " Automated tracking of structural steel members at the construction site", ISARC 2000, 1201-1206.

Jaselskis, E. J., Anderson M. R., Jahren, C. T., Rodriguez, Y., Njos S.: "Radio-frequency identification applications in construction industry", J. of Const. Eng. and Manag., 1995, Vol. 121, No. 2, 189-196.

Marconi InfoChain: http://www.marconi-infochain.com/tech/rfid.htm, 06/27/2002

Paolo, M. D., Furr, W., Hearn, N. N: "RFID: Replacement or supplement to barcodes?", http://citeseer.nj.nec.com/490581.html, 06/27/2002

Pheng, L.S., Chuan, C. J.: "Just-In-Time management of precast concrete components", J. of Const. Eng. and Manag., 2001, Vol. 127, No. 6, 494 –501.

Schell, D.: "Oil refinery uses RFID to eliminate rework and reduce data entry by 70%", Business Solutions, January 2001.

Texas Instruments, http://www.ti.com/tiris/docs/manuals/pdfSpecs/RI-I02-112A.pdf, 06/27/2002.

FIGURES

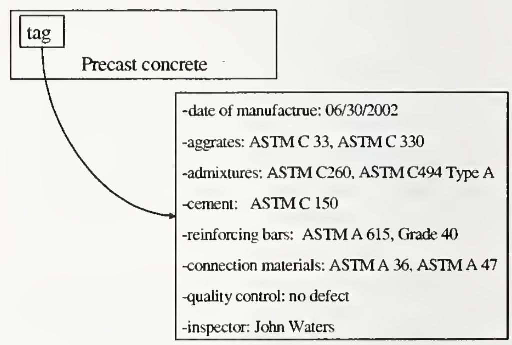


Figure 1: Information entered to a tag during casting of a precast member

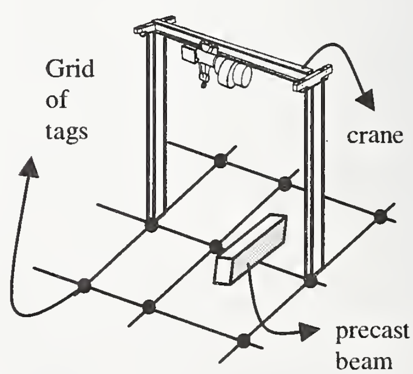


Figure 2: Grid of transponders in the storage area

MEMS Application in Pavement Condition Monitoring – Challenges

by

Nii O. Attoh-Okine*
Stephen Mensah[#]

ABSTRACT

Although the universal definition of MEMS product possesses a number of distinctive features, they are miniature systems involving one or more micro-mechanical components or structure. Infrastructure assessment has recently addressed an important field for MEMS application. For example, satisfactory laboratory experiment in the area of concrete monitoring has been developed. Although the experiment is at an infant stage, it demonstrates the enormous potential of MEMS application in civil infrastructure systems. This paper proposes to develop a framework for the application of MEMS technology in pavement condition monitoring and evaluation. Some difficulties which may be encountered, especially in asphaltic medium and loading condition of the pavement will be discussed.

Keywords: devices, evaluation, microelectromechanical, monitoring, non-destructive, pavement, system

1. INTRODUCTION

Pavement condition monitoring and evaluation is important in many areas of pavement engineering, especially in pavement management. The in-service performance of the pavement depends on consistent, cost-effective and accurate monitoring of condition for early scheduling of repair and maintenance. For the past decade, nondestructive evaluation (NDE) testing has played a major role in pavement condition monitoring, assessments and evaluation. The NDE evaluation methods include seismic evaluation such as wave propagation, vibration methods, acoustic and ultrasonic methods, electromagnetic method and electrical resistivity methods.

The accuracy and repeatability of the interpretation of these NDE techniques are dependent on the limitations imposed by the operating conditions and material properties. For example, the use of acoustic methods is well established in water but in a dense material, especially concrete pavements, specially developed sensors are required [1].

Researchers from diverse disciplines especially electrical, computer and mechanical engineering, have been drawn into vigorous efforts to develop sensing technologies and nano-technology in infrastructure condition assessment [2], crack detection [3] and building monitoring [4]. Recently, MEMS have been proposed as a tool in infrastructure condition monitoring. Researchers over the years have developed microfabricated sensors for measuring position, acceleration, pressure, force, torque, flow, magnetic field, temperature, gas composition, himidity, and biological gas/liquid molecular concentration [5]. Its miniature size has enabled widespread use in the medical field and the automotive industry. MEMS have had a tremendous impact on our modern society: it has led to creation of jobs; it has a significant leverage factor enabling the production of sensor-based systems exceeding sensor cost by several orders of magnitude; and in some instances have even become critical components that determine the feasibility of new products [6]. While the success of MEMS continues to grow, it appears its use in civil infrastructure monitoring is yet to make any impact.

This paper delves into the potential application of MEMS in infrastructure monitoring and some of the technological as well as technical issues to grapple with.

2. MICROELECTROMECHANICAL SYSTEMS (MEMS)

- MEMS can generally be characterized by the following. They are miniature systems involving one or more micro-machined components or structures. The inherently small size of MEMS enables high-level functions by virtue of the large-scale integration process. This permits multiple functions or capabilities integrated o the same silicon chip or package for greater utility.

While MEMS devices are sought to be of somewhat dimensions of only a few micrometers, the small size of a device does not automatically qualify it as a MEMS device. The distinctive features are [8]: miniaturization, micro-electronic and multiplicity. Also he device consists of components which act together in a systematic fashion for a desired purpose. MEMS has been described as a toolbox, a physical product and a methodology [7]:

- It is a portfolio of techniques and processes to design and create miniature systems;

It is a physical product often specialized and unique to a final application

The MEMS technology will have an impact on engineering in the following ways [9]:

- By causing orders of magnitude increase in the number of sensors and actuators
- By enabling the use of very large-scale integration (VLSI) as a design and synthesis approach for electromagnetics.
- By becoming a driver for multiple, mixed and emerging technology integration.
- By being both a beneficiary of and a driver for information systems.

In addition to the potential economic benefits MEMS has the ability to integrate mechanical (or chemical, biological and environmental) functions. It will also allow for consideration of concepts such as the highly distributed networks for the condition monitoring of large civil infrastructure systems. The fundamental technological issues in MEMS include:

- Materials
- Machining process
- Micro-mechanical devices
- Application

2.1 Materials

Silicon has been used extensively however there have been advances in new materials exploration and the study of material properties for microsensors and actuators has opened up new frontiers of materials. The progress of microactuator technology depends critically upon the development of actuation forms that are compatible with the materials and processing technologies of silicon microelectronics [5]. Mechanically, silicon is an elastic and robust material whose characteristics have been well studied and documented. Furthermore silicon as a material exists in three forms: crystalline, polycrystalline or amorphous. Each of the forms has unique properties which make them useful for different applications. Another important property of silicon is its ability to integrate with electronic circuits and sensors. Three important material properties which are of great interest to MEMS are:

- Piezoresistivity – the phenomenon where a material's resistivity changes with mechanical strain.
- Piezoelectricity – when crystals exhibit the peculiar property of producing an electric field when subjected to external force.
- Thermoelectricity – the interaction between electricity and temperature.

The choice of material depends on its compatibility with the current fabrication process as well as other materials to be used in the device fabrication. A material property of significance cannot be exploited if this compatibility is not met.

2.2 Micromachining

Three fabrication routes or methods account for the majority of MEMS devices: surface micromachining, bulk micromachining and molding. Surface micromachining [10] uses the CMOS (complementary metal oxide semiconductor) process to fabricate VLSI (very large scale integration) devices. In the surface micromachined MEMS, the layers are patterned and etched to yield electromechanical elements to allow motion of the mechanical layers. Bulk micromachining involves etching features directly into the silicon wafers, it is an important consideration in MEMS where higher mechanical power or force levels are desired [11]. Molding is the third prevalent manufacturing process used for MEMS. It is the creation of the mechanical elements of the device by the deposition of material into a microfabricated model.

2.3 Microelectromechanical Devices

Microelectromechanical devices generally comprise of microsensors, microactuators, and electronic circuits or signal processing units integrated on a single chip. The miniaturization of the devices makes them inherently smaller, lighter and faster than their macroscopic counterparts and are usually more precise. The microsensors sense or collect information from the environment whereas microactuators alter the state of the environment. This process essentially involves the conversion of a mechanical force or physical effect in the environment (such as change in pressure or temperature) into electrical signals and vice-versa. This is possible by taking advantage of peculiar material properties such as piezoresistive, piezoelectric and thermoelectric effects to convert the change in the environment to electricity. The signal processing unit or electronics will then process the information and provide inputs to the actuators. The actuators in turn can manipulate the environment for a desired purpose or trigger another action to compensate for the change. Microactuators are the key devices allowing MEMS to perform physical functions. They are categorized in two perspectives: one based on driving forces and the other based on mechanisms.

2.4 Application

Microdevices are being embedded in structures to enhance their performance. Such structures are called smart structures or smart materials. Smart materials and structures technology is a new field of study that is finding its way into many applications in civil infrastructure systems. The applications include structural control, condition health monitoring, damage assessment, structural repair and integrity assessment. More recently and extension has been made to cover asset management and preservation and operation of civil infrastructure. The drawback to extensive MEMS application is that there are no generic MEMS products but rather they are application specific. The vast majority of applications require unique solutions that often necessitate the funding a completion of an evaluation or development program. This process could take several years, typically 2-5years, and it impacts directly on the performance of MEMS devices on the market [5].

3. PAVEMENT CONDITION MONITORING

3.1 Infrastructure maintenance is undertaken based on the perceived health of the infrastructure based on information gathered from the performance indicators. The need to collect data reflective of the true state of health of the infrastructure is therefore crucial for an efficient management of the system. Condition based maintenance (CBM) or on-demand maintenance are gaining currency due to the superior advantage they afford. Microsensors can be embedded unobtrusively and inconspicuously in structures to monitor parameters critical to the safe operation and performance. Bennet et al. [12] developed a wireless monitoring system for highways. The system consists of a retrofitted instrumented asphalt core, which is bonded into the pavement structure. The core contains all of the electronics necessary to record two pavement temperatures (surface and base) and two strains (longitudinal and transverse). The data when collected is transmitted via low power radio link to a receiver and data logger positioned by the side of the road. Also Sackin [13] proposed a new laboratory approach for the feasibility of embedded microdevices for infrastructure monitoring in concrete. This microdevice will be termed as "smart aggregate". One of the objectives of the work is to investigate how an embedded device will behave under real working and environmental conditions. Although the laboratory test was successful, most questions vital to the MEMS application in infrastructure systems were not adequately considered. Some of these will be highlighted in the subsequent paragraph. Table 1 summarizes the potential application of MEMS based on condition deterioration mechanism.

3.2 Important Considerations

It is envisaged that MEMS devices can be embedded in the road infrastructure to monitor the condition at all times. The information collected can then be relayed via wireless technology into a database. Or MEMS devices with remote query capability can be interrogated at any given time to assess the condition of the pavement system as mentioned earlier. For the successful application and implementation of MEMS in pavement for continuous monitoring, some issues need to be addressed. Some of these issues are listed below:

- The effect of asphaltic medium on MEMS.
- How many MEMS devices are to be installed per 1km of pavement
- Where will the designer install or embed the microdevices (MEMS) in the pavement?
- Pavements have a life span of 20-30 years, will MEMS devices be able to perform throughout this period?
- How will environmental conditions affect the performance of MEMS in pavement condition monitoring?
- How will the effects of chemical medium such as corrosion in reinforced concrete structures, affect the performance of MEMS.
- Although the cost of MEMS is predicted to be reasonable, will it be a cost-effective method of collecting continuous data?

Some of these considerations have been addressed in other fields. For example, MEMS has been used in various chemical media in bioengineering applications [14].

4. CONCLUSIONS

The paper presents some challenges that need to be addressed for a successful implementation of MEMS technology in pavement monitoring. For example, laboratory investigation on the behavior of embedded MEMS in asphaltic material medium under dynamic loading needs to be addressed. Finally, it is required to test and track the field behavior of the pavement to establish both repeatability and long-term behavior of MEMS embedded in the pavement.

5. REFERENCES

1. W. Uddin and F. Najafi, "Deterioration mechanisms and nondestructive evaluation for infrastructure life-cycle analysis". Infrastructure Condition Assessment: Art, Science and Practice. Edited by Mitsuru Saito, pp 524-533, 1997.
2. R. Ansari. "Theory and applications of integrated fiber optic sensors in structures". Intelligent Civil Engineering Materials and Structure 22, pp 2-28, 1997.
3. P. Rossi and F. LeMaou. "New method for detecting cracks in concrete using fiber optics". RILEM, Materials and Structure 22, pp. 437-442, 1989.
4. P. L. Fuhr et al. "Performance and health monitoring of the Stafford Medical building using embedded sensors". Journal of Smart Materials and Structures 1, pp. 63-68, 1992.
5. M. Mehregany. Microelectromechanical Systems". IEEE Circuits and Devices 9 (4), pp 14-22.
6. J. Bryzek. "Impact of MEMS technology on society". Sensors and Actuators A (56), pp 1-9.

7. N. Maluf. "An introduction to microelectromechanical . systems engineering". Artech House Publishers, 2000.
8. H. Fujita. "Microactuators and Micromachines". Proceedings of IEEE, 86 No. 8, pp 1721-1732, August 1998.
9. K. J. Gabriel, Microelectromechanical Systems". Proceedings of IEEE, 86 No. 8 pp. 1534-1535, August 1998.
10. J. M. Bustillo, R. T. Howe, and R. S. Muller, "Surface micromachining for microelectromechanical sytems". Proceedings of IEEE, 86 No. 8 pp 1552-1574, August 1998.
11. S. M. Spearing. "Materials issues in microelectromechanical systems (MEMS)". ACTA Material 48, pp. 179-196, 2000.
12. R. Bennet. "Wireless monitoring of highways". SPIE Conference on Smart Systems for Bridges, Structures and Highways 3671, pp 273-182, 1999.
13. M. Sackin. "A feasibility study of embedded microdevices for infrastructure monitoring: Smart aggregate for concrete". A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science, December 1999.
14. A. Pedrochhi et al. "Perspectives on MEMS in bioengineering: a novel capacitive position microsensor". IEEE Transactions on Biomedical Engineering 47 No. 1 pp 8-11, January 2000.

Table 1. Potential application of MEMS based on condition deterioration mechanism.

Condition deterioration	Deterioration mechanisms		Potential use of MEMS
	Load/usage, environment, material degradation, construction quality, interaction	Other	
1. Surface defects	Primary (environment, material degradation) Secondary (load/usage, poor construction quality) Interaction (material/environment and extended by load)	Man-made defects, maintenance patches	Fair
2. Deformation	Primary (environment, material degradation) Secondary (load/usage, poor construction quality) Interaction (material/environment/load)	Loss of support, inadequate maintenance	Excellent
3. Cracking, disintegration	Primary (environment, material degradation) Secondary (load/usage, poor construction quality) Interaction (load /environment/ material degradation)	Accidents, inadequate maintenance	Very good to excellent
4. Failure (a) Aging/inadequate structural capacity or retirement	The facility is structurally deficient because the limiting threshold values of (1) surface defects (2) deformation (3) cracking and disintegration are exceeded or the facility is retired because it is functionally obsolescent. Interaction (nature)	Capacity and safety considerations, inadequate maintenance	Good
(b) Catastrophic failure	Primary causes: earthquake, floods, freeze/snow, ice, tornado/cyclones, wind storms, foundation and soil failure and sink holes. Interaction (poor construction quality/design deficiency	Fire, arson, terrorist act or other accident	Poor

Towards the Ultimate Construction Site Sensor[1]

by

William C. Stone[2] and Maris Juberts[3]

ABSTRACT: The NIST Construction Metrology and Automation Group (CMAG), in cooperation with the NIST Intelligent Systems Division (ISD), is developing performance metrics and researching issues related to the design and development of a "Next Generation LADAR" sensor that will enable general automation in structured and unstructured environments. This paper quickly reviews the basic physics and implementation of various LADAR technologies, describes the problems associated with available "off-the-shelf" LADAR systems, summarizes State-of-the-Art work underway around the world, and elaborates on general directions that advanced research in this area of sensor design will take in the coming years and its likely impact on construction automation.

KEYWORDS: AM-modulation, construction automation, FM-modulation, laser radar, multiple returns, phantom points, pulse time-of-flight, range image sensors

1.0 INTRODUCTION

Laser Detection and Ranging (LADAR) is currently poised to become the ubiquitous 3D spatial measurement tool in many disciplines. Initially used for remote sensing and aerial surveying, LADAR applications now include reverse engineering (3D models), ground surveys, automated process control, target recognition, and autonomous machinery guidance and collision avoidance to name just a few. Efforts are currently underway at NIST to develop national artifact-traceable LADAR calibration facilities; to develop rapid, LADAR-based long range autoID systems; and to establish the scientific and engineering underpinning needed to develop miniature, high resolution next-generation LADAR systems.

The power of LADAR lies in the inherent 3D nature of the data it produces, namely spatial coordinates associated with each pixel in a so-called "range image" acquired by the device. A range image is effectively a spherically acquired (r, θ, φ) dataset mapped to a 2D matrix, or "frame." LADAR frames are frequently presented as false color depth images. Additional data, including reflectance intensity associated with each pixel and multi-spectral intensity information, are commonly available. Color reflectance intensity (as opposed to active illumination frequency-specific reflected intensity) is obtained from co-boresighted RGB CCD sensors. Such a wealth of information can be rapidly segmented for use by a wide variety of real-time systems for machine control and post-processed for such metrology applications as as-built geometry checking for buildings and other civil infrastructure. This said, why are we not seeing LADAR systems on every construction site? The reasons most frequently cited are: slow speed of operation, bulky, high cost, and widely varying accuracy that presently lacks standardized calibration metrics. There are other related issues such as

[1] Official contribution of the National Institute of Standards and Technology (NIST); not subject to copyright in the United States.

[2] NIST, BFRL, Construction Metrology and Automation Group, Mail Stop 8611, Gaithersburg, MD 20899-8611; william.stone@nist.gov.

[3] NIST, MEL, Systems Integration Group, Mail Stop 8230, Gaithersburg, MD 20899-8230; maris.juberts@nist.gov

methods for processing the data (both real-time and offline) but that is the subject of another paper. Frequently the performance of a LADAR is defined by the following metrics:

- Maximum permissible illumination power
- Sensor horizontal Field of View (FOV)
- Sensor vertical FOV
- Wavelength of optical source
- Maximum distance to be measured
- Measurement time
- Measurement resolution (depth)
- Measurement resolution (angular)
- Range Measurement accuracy

To these one frequently must consider:

- Intensity of background (passive) illumination
- Color temperature of the (passive) background
- Target reflectivity (texture, color, specularity)
- Angle of beam incidence on the object
- Overall size (volume) of the sensor
- Manufactured cost of the sensor

At the conclusion of this paper we will present a set of design criteria we feel are representative of those needed to achieve ubiquitous use of LADAR sensing for construction operations. We will also comment on the research needed to achieve a physical sensor meeting such criteria.

2.0 LADAR PHYSICS SUMMARY

2.1 Pulse Time-of-Flight (TOF)

Figure 1 shows a "family tree" of LADAR devices that have at one time or another been built to operate at optical and near-optical wavelengths. References 1 and 2 provide useful in-depth discussions on many of these devices. The simplest of the concepts uses pulse time-of-flight (TOF), as illustrated in Figure 2. An illumination pulse is generated, frequently by means of a Nd-YAG microchip laser, and the time of this event is made available to a timing circuit. The beam traverses a distance equal to *2d* and arrives at a photonic detector in time:

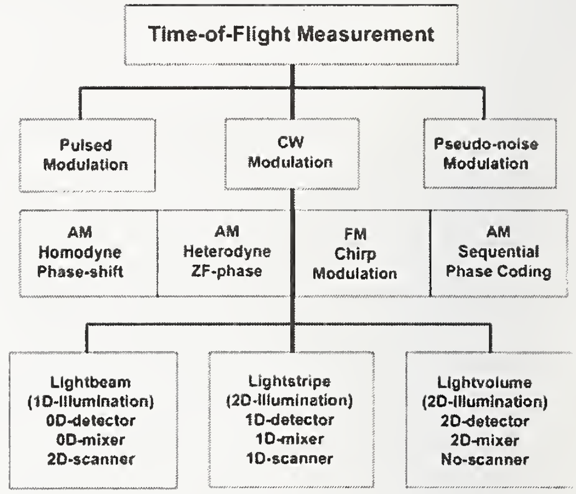


Figure 1: "Family Tree" of optical and near-optical wavelength time-of-flight range measurement devices.

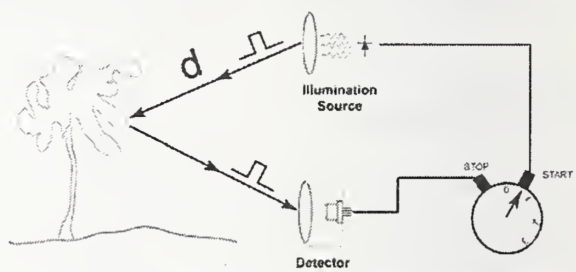


Figure 2: Fundamental physics of pure "pulsed" time-of-flight. Key performance metrics are synchronization precision of pulse initiation between the source and detector, pulse width, and detector bandwidth.

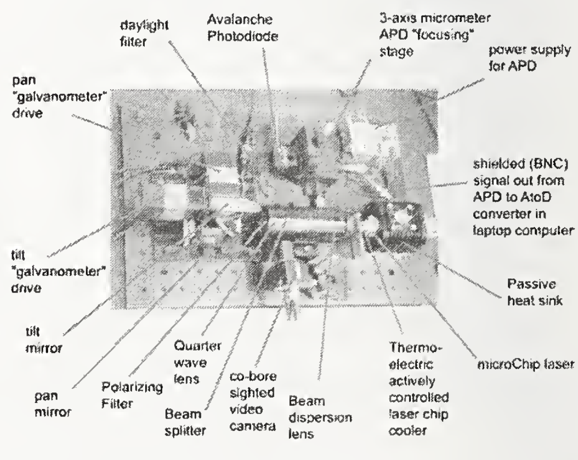


Figure 3: Typical physical implementation of a pure pulsed time-of-flight LADAR Scale is approximately 300 mm long by 150 mm wide.

$$t = \frac{2d}{c}$$

where c = speed of light. A clock time of 1 ns represents a 300 mm round-trip flight or an absolute range of 150 mm. One can immediately see that an extremely accurate clock is required to achieve a level of accuracy sufficient for autonomous fabrication. Figure 3 shows a typical physical implementation of a pulse TOF LADAR [8]. The microchip laser generates 1 ns pulses at a rate of 10 kHz, producing an unambiguous range interval of 15 km. The signals are detected by an avalanche photodiode (APD). In this particular implementation an optical beam splitter is used to divert a portion of the source signal to the APD, thus providing the "start timer"mark for range determination. The accuracy of such systems depends on a number of factors including the pulse width, detector electronics bandwidth, and the processor implementation. If the APD bandwidth is 2 GHz and the analog-to-digital readout (also known as ROIC) is matched, then the timing "bin" width is *approximately* 0.5 ns, but since that is round-trip, the range bin accuracy is thus 75 mm. This "0-D" ranging system is then *scanned* in 2D using electromechanically steered mirror systems.

2.1.1 Limitations of Pulse TOF

Thus far we have made the assumption that all of the photons that are generated hit one specific object and are reflected back to the detector in a narrowly discriminating beam yielding one range measurement per pixel. This is not the case as shown in Figure 4. Due to imperfect optics and atmospheric dispersion the source illumination beam (pulse) expands with range; good industrial LADARS that have achieved near-diffraction limited optics have beam dispersion angles of around 0.2 mrad. Even so, this produces a finite beam diameter at 100 m of around 20 mm. This has interesting physical consequences. Because the un-ambiguous range of the device shown in Figure 3 is on the order of 15 km, one receives, in time, responses from photons from the illumination pulse arriving at different times related to

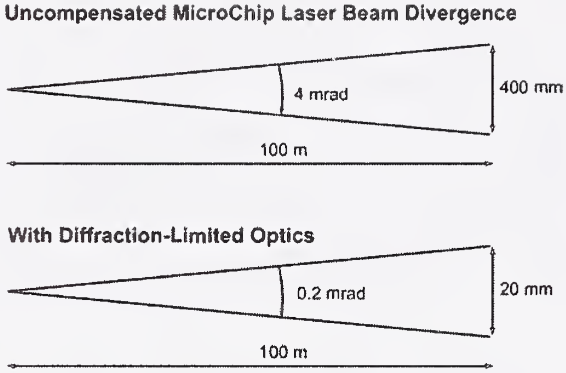


Figure 4: Source beam divergence variance and its effect on absolute beam diameter at 100 m range.

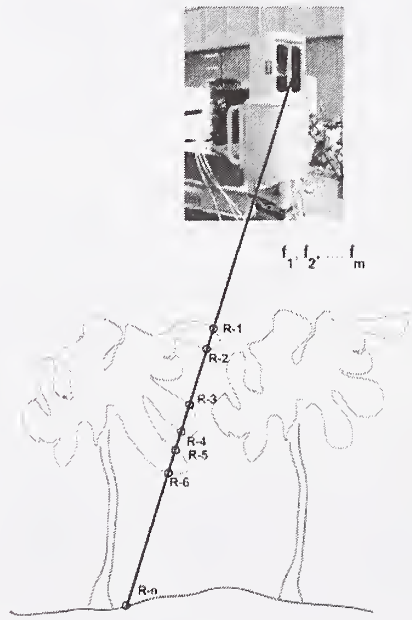


Figure 5: Due to beam divergence photons associated with a single "pixel" in a LADAR frame may represent significantly differing range data.

the different objects they hit within the cone of the dispersed beam. Figure 5 illustrates this point. A single "pixel" in the LADAR frame will in fact have multiple valid ranges. Figure 6 shows a time-domain response out to a range of approximately 120 m. Any strong return above the noise threshold represents a valid object detection. Thus, one could store not one value, but a vector of values, for each pixel. Some LADARs now being developed illuminate at more than one wavelength. The response at each

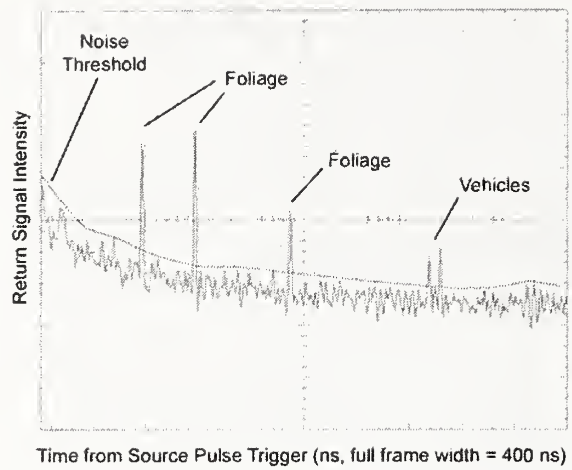


Figure 6: Full time-domain response for a 1ns pulse. Response was captured in 801 "range bins." The time-width of each bin was 0.5 ns, or 150 mm.

pixel can then be represented by a matrix with n returns per source frequency f:

$$\begin{bmatrix} R_{1,f_1} & R_{1,f_2} & R_{1,f_3} & \cdots & R_{1,f_m} \\ R_{2,f_1} & \cdots & \cdots & \cdots & \cdots \\ R_{3,f_1} & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{n,f_1} & \cdots & \cdots & \cdots & R_{n,f_m} \end{bmatrix}$$

At present, no commercial pulse TOF LADAR provides the user with this kind of pixel response matrix (or even a single frequency time domain vector). Instead, it is common to average earlier arrivals that have strong S/N (signal-to-noise) ratios and report that as a single range per pixel. The results of this averaging are shown in Figure 7, where non-existent "phantom" points become part of the point cloud data set. The point here is that each of the valid returns shown in Figure 6 represent usable engineering information that is presently not available. Pulse TOF systems are limited in their accuracy not only by the bandwidth of the detector (currently pegged at around 2 GHz, although current research in fiber optic tele-communications is pushing this higher), but also by the pulse width of the source illumination, since edge detection is enhanced by a shorter, sharper pulse. The shortest pulse source cur-

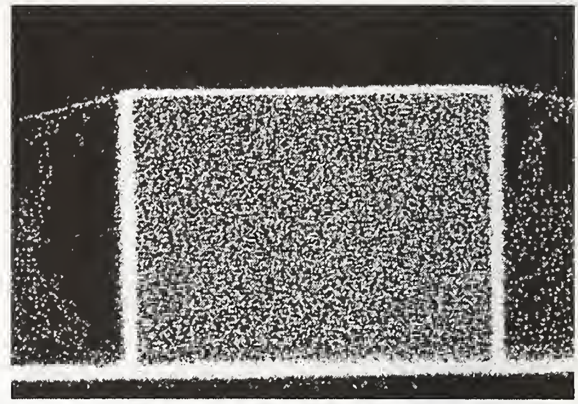


Figure 7: "Phantom" points (sloping ledges leading diagonally downward from the top) generated during scans of a rectangular box using single-point-per-pixel reporting from an industrial pulse TOF LADAR. Averaging of multiple returns (see Figure 6) to provide a single range per pixel leads to the erroneous reported locations.

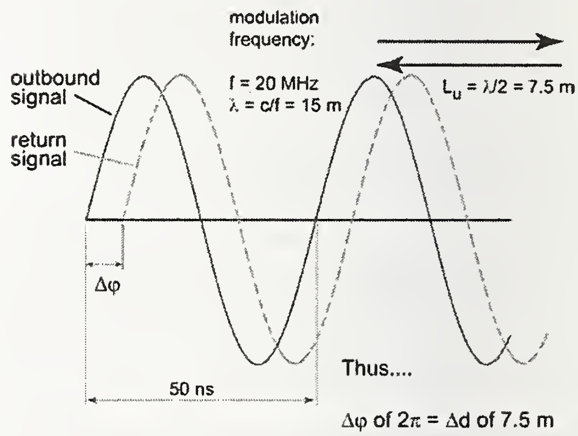


Figure 8: Phase-based determination of time-of-flight. Primary modulation frequency establishes the maximum unambiguous range for the device.

rently used in LADAR devices is 250 ps [5].

2.2 AM Homodyne Phase Modulation

One way to beat the "brute force" approach to time-of-flight accuracy is through the use of phase detection. This concept is illustrated in Figure 8. If the source is modulated at a single sinusoidal frequency, f, then a phase shift of

$$\Delta\phi = 2\pi f T = 2\pi f \frac{L}{c}$$

will be observed between the transmitted and received signal. Since the range, R, is equal to half the round-trip distance, L,

$$R = \frac{\Delta\phi c}{4\pi f}$$

The unambiguous range resolution is directly proportional to the source modulation frequency, f, while the accuracy is directly proportional to the signal-to-noise ratio. The two are directly related; the tighter the unambiguous range the finer can be parsed the phase difference, thus improving the accuracy. For the case shown in Figure 8, the unambiguous range is half the wavelength at f= 20 MHz, or 7.5 m. Signal-to-noise ratio can be improved by integration over many cycles, but at the cost of raw throughput (frame rate).

Range determination using phase measurement is adapted from earlier work in radar in which the source and received signals are mixed and the phase, amplitude, and offset of the resulting signal are determined through Fourier theory. If the resulting signal is sampled at four intervals separated by phase angles of π/2, it can be shown [2,6] that the phaseshift, Δϕ, is given by

$$\Delta\phi = \arctan\left[-\frac{A_1 - A_3}{A_0 - A_2}\right]$$

where A_0, A_1, A_2, A_3 represent the integration of the mixed signal over the intervals shown in Figure 8. The trick is in the formation of the

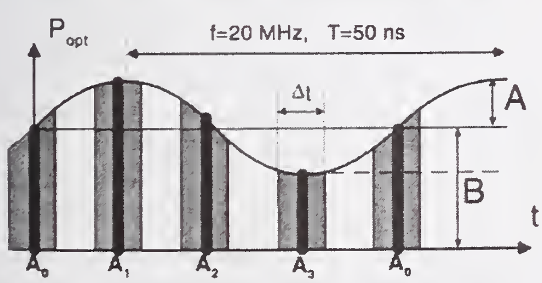


Figure 8: In a phase-based solution the transmitted and received signals are mixed. Discrete Fourier Transform (DFT) theory allows extraction of the phase, amplitude, and offset by sampling four points at an interval of π/2 along the resulting waveform.

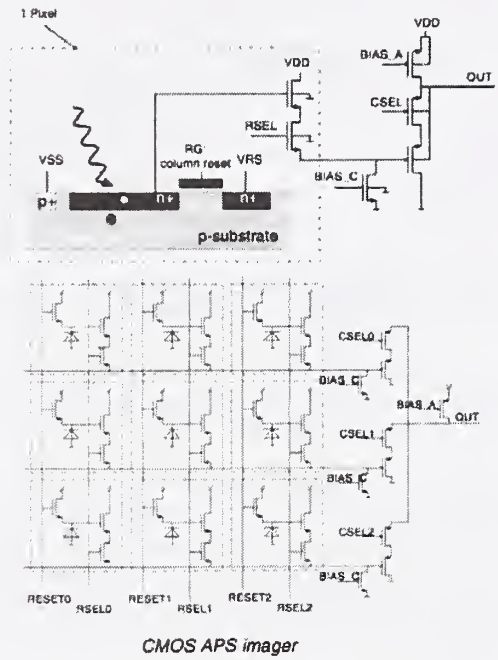


Figure 9: Example of an "Active Pixel Sensor," one of many versions of a Focal Plane Array (FPA) that can be used to simultaneously sample hundreds to thousands of pixels. FPAs form the core of "flash" LADAR and can utilize either phase or pure time-of-flight ranging.

mixed signal and in the integrations to be performed. One of the more clever solutions to this problem makes use of so-called *photonic* mixing (Figure 9) in which a standard CMOS photo diode is reverse biased by the same frequency source that modulates the transmitted illumination signal. This diode responds directly to the mixed signal consisting of the modulated bias and the incoming photo electrons. Timing circuitry dumps the white portions of the curve in Figure 8 while summing (through a capacitor) the regions A_0, A_1, A_2, A_3 and selectively storing those values. It is possible to implement such detectors in 2D arrays in the form of *Focal Plane Arrays* (FPAs) as shown in Figure 9. Individual pixels within a LADAR frame are mapped (optically) to pixels on the FPA. Similar FPAs can be constructed to work on the pulse TOF principle. Both approaches are referred to as *flash* LADAR.

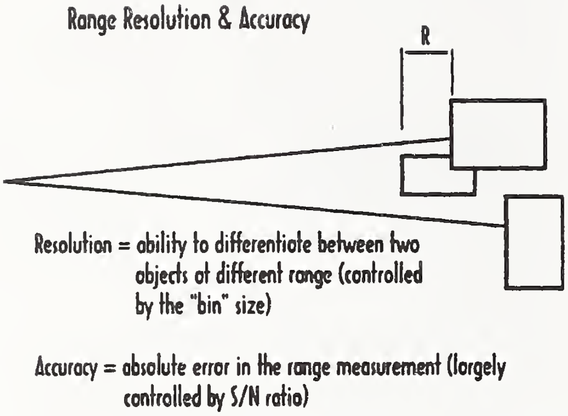


Figure 10: Due to beam divergence and varying reflectivity of target surfaces, both accuracy and resolution are affected for all types of LADAR, but particularly for AM-homodyne class devices which in effect integrate the reflected photons from all surfaces that are illuminated, thereby producing an erroneous average range.

2.2.1 Limitations of AM Phase Measurements

Pure homodyne (single frequency) LADARs suffer from two significant limitations. As the modulation frequency is increased, thereby improving potential accuracy, the un-ambiguous range is reduced -- leading to *aliasing* or false targets. A solution to this problem is to use multiple frequencies [4] in which a lower frequency signal is used to establish an un-ambiguous interval over a longer distance within which the higher frequency response is located. The mathematics for solving this approach are not significantly more complicated than for the pure homodyne solution, but it has yet to be implemented in silicon. A more serious problem with phase-based measurement is illustrated in Figure 10. For the same reasons that multiple returns are received in pulse TOF systems, the integration response of the photonically mixed signal includes the energy reflected from all the surfaces shown in Figure 10, regardless of their different ranges, as long as they are within the illumination cone of the dispersed source beam. Thus, the range reported is an average of the objects within that pixel, leading to the same type of *phantom points* described earlier.

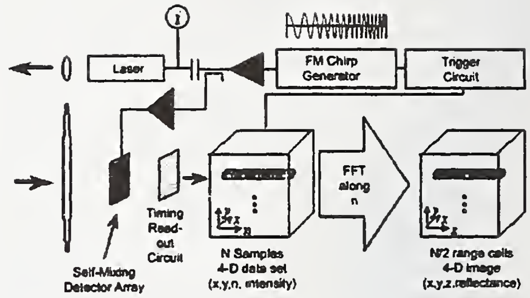


Figure 11: Block diagram for an experimental FM-CW LADAR. This design overcomes range averaging issues with AM-CW system [7].

2.3 FM-CW Modulation

It is possible to recover multiple ranges per pixel in a frequency modulated LADAR by producing a very specifically-shaped source pulse, known as a chirp. The theory for such an approach has been around since the advent of radar [7] but only recently has been adapted to optical systems. In a *chirp* pulse the frequency is varied linearly with time; typically over the range of 200 MHz to 600 MHz with a frequency response of 800 MHz or more. The range resolution (per pixel) in such a device is given as

$$\Delta R = \frac{c}{2\Delta f}$$

in which Δf is the chirp frequency step size. The detector in Figure 11 consists of a photonic FPA mixer and CMOS read out circuit. The result is a frequency-domain signal (an integrated intensity for each of n frequencies contained in the chirp bandwidth). This signal must be subsequently processed, using an FFT, to obtain a time domain response. Nyquist sampling criteria increases the range bin size by a factor of 2.

The chief limitations of the FM-CW LADAR approach are the need for additional front end hardware (complicated chirp-generation electronics that add both cost and an additional level of noise), the frequency response of the laser source (which may be bandwidth limited); and the manifold increased mathematical post-processing that is involved.

3.0 IDEAL LADAR PERFORMANCE

NIST has for a number of years conducted research in autonomous robotic platforms and machinery and has reached a consensus regarding the required performance of a “vision” system needed for effective control. A LADAR meeting these criteria would have the following attributes:

Illumination Source: eyesafe
Field of View (FOV): 60° x 60°
Range resolution: 1 mm @ < 15 m
3 mm @ < 100 m
Angular resolution: < 0.03°
Frame rate: >10 Hz
Size: coffee cup
Cost: < $1000 US

Some of these criteria can be met by existing systems but most cannot. Frequently there is a tradeoff between speed and accuracy. 2D frames can be created using high resolution laser rangers by a method known as "scanning" in which a single-degree-of-freedom laser radar is mechanically swept over the scene using either encoder-equipped pan/tilt servos or a rotating mirror combined with either a pan or tilt servo. Because of this mechanical reliance, however, these systems have inherent speed and accuracy limitations associated with the servos and their encoders. New work in FPA design shows promise for both improving range resolution as well as speed. And only this latter approach shows promise for the miniaturization needed to achieve the last two criteria listed above. The FOV and angular resolutions listed above translate to a 2048 x 2048 pixel FPA. The largest range-imaging FPA yet to be demonstrated successfully is 25 x 64, although several labs have 128x128 arrays under development. There are other significant issues: to illuminate a large FOV requires considerable laser power, possibly making the output unsafe (eye-safe) at most of the compatible wavelengths. Parallel FPAs or MEMs-based steering provide possible solutions.

4.0 Conclusions and Avenues for Research

The joint NIST effort to develop a Next Generation LADAR (NGL) for autonomous machine control and construction metrology has identified four key areas for research needed to achieve a functioning sensor with the above specifications:

• *Ultra-fast Chip-Level Laser technology*: build and test compact pulsed femtosecond (~10^{-13} s) coherent laser sources. The performance, power needed, source-detector cross-talk and ranging accuracy of prototype lasers and APD receivers will be characterized at NIST. It will be determined whether nonlinear quantum dot saturable absorbers can be implemented in micro-laser cavities to achieve shortest possible pulse duration. Initial efforts will be to experimentally determine if extensions to known micro-chip laser systems can achieve desired femtosecond performance followed by work on candidate new designs.

• *Ultra-precise Chip-Level Time and Frequency Standards*: Investigate development of MEMs/CMOS-based manufacture of on-chip, high accuracy time and frequency standards and phase correlators. Investigate development of candidate designs for CMOS implementation of the best timing and frequency designs and ultimately to the fabrication of test articles. Refine designs of an on-chip timing and frequency standard that could be readily integrated with a candidate APD/FPA for testing with either the chip light source developed in task 1 or by means of external femtosecond laser sources.

• *Fast Beam Steering*: There are several different technologies available for fast beam steering. The most popular of these are acousto-optic, tilting mirrors and electro-holography. Miniature beam steering devices are being investigated at NIST based on these technologies. The tilting mirror technology has received special attention recently, because MEMS-size mirror arrays have been built. Acousto-optic devices use acoustic waves propagating in a variety of optic materials to control the refractive index of the material and thus the angle of the output light beam. The electro-holography technology involves the writing of Bragg grating holograms of specific wavelengths on photo-refractive crystals. NIST is evaluating these beam steering technologies and will select the most promising of them for prototype devices.

• *Systems Integration and Performance Analysis*: Early analysis here has focused on the advantages and disadvantages of direct TOF versus phase correlation methods for range determination within the context of systems capable of being fabricated using MEMs and CMOS technology combined with InGaAs / InP detectors that are either bump or bridge-bonded to the read out circuit. This effort is also investigating issues and techniques for integrating a single channel micro-steerable LADAR. Computer simulations are being developed to evaluate the feasibility of miniaturization, operation in various environments, sensor detection limits, and performance characteristics - distance, accuracy, and speed.

The results of this research in the coming years will prove the feasibility of building the ultimate construction site sensor.

5.0 REFERENCES

1. **Besl, P.J.,** "Active, Optical Range Imaging Sensors," Machine Vision and Applications (1988) 1:127-152.

2. **Lange, R.**, "3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology," Doctoral Dissertation, Dept. of Electrical Engineering and Computer Science at University of Siegen, Siegen, Germany, June 28, 2000.

3. **Stettner, R., and Bailey, H.W.**, "3D Imaging Laser Radar," United States Patent 6,133,989, October 17, 2000.

4. **Frohlich, C., Mettenleiter, M., Hartl, F., Dalton, G., and Hines, D.**, "Imaging Laser Radar for 3-D Modeling of Real World Environments," Sensor Review, Volume 20, No. 4, 2000, pp 273-281.

5. **Aull, B.F. et al**, "Geiger-Mode Avalanche Photodiodes for Three-Dimensional Imaging," Lincoln Laboratory Journal, Volume 13, Number 1, 2001.

6. **Spirig, T.,** "Smart CCD/CMOS Based Image Sensors with Programmable, Real-Time, Temporal and Spatial Convolution Capabilities for Applications in Machine Vision and Optical Metrology," Doctoral Dissertation, Swiss Federal Institute of Technology, Zurich, 1997.

7. **Stann, B.L., Ruff, W.C., and Sztankay, Z.G.,** "Intensity-Modulated Diode Laser Radar using Frequency-Modulation / Continuous-Wave Ranging Techniques," Optical Engineering, November 1996.

8. **Schilling, B.W., Barr, D.N., Templeton, G.C., Mizerka, L.J., and Trussell, C.W.**, Multiple-return laser radar for three-dimensional imaging through obscurations, APPLIED OPTICS, Vol. 41, No. 15 pp 2791-2799, May 2002

SESSION 8

CONTROL SYSTEMS FOR CONSTRUCTION EQUIPMENT

A COST-EFFECTIVE POSITIONING SOLUTION FOR ASPHALT ROLLERS BASED ON LOW-COST DGPS RECEIVERS

by

Jaroslaw Jurasz[1], Karl Ludwig Kley[1]

ABSTRACT: As of today, the DGPS-based Computer Integrated Road Construction systems for compaction support require high investments (about 50% of total machine cost), mostly due to the high cost of the positioning equipment. Worksite tests show that the majority of the compaction errors are serious omissions of the compaction plan. The article discusses the accuracy requirements and implementation structure of a cost-effective system based on the low-cost DGPS receivers, aimed at detection and correction of the major compaction errors. First results of the experimental validation confirm the feasibility of the approach. The full-scale implementation is currently under development and scheduled for worksite tests. The availability of cost-effective and robust positioning solutions should improve the acceptance of the DGPS-based compaction support systems.

KEYWORDS: compactor pass map; computer integrated road construction; low cost DGPS; positioning; unscented Kalman filter

1. INTRODUCTION

The purpose of the Computer Integrated Road Construction (CIRC) is to increase the quality of the road works by tracking the construction processes in the real time [1]. As the basic concept is based on the comparison of the actual position of the machine's tool with the desired value, e.g. with a digital terrain model (DTM), the positioning sub-system is a crucial element of every CIRC system and greatly influences its performance.

As of today, the CIRC systems use two basic positioning techniques: the Robotic Total Station for high precision measurements (mm range) and Differential GPS (DGPS) receivers for lower accuracy (cm-dm range). The DGPS-based compaction support systems introduced recently, both in research and commercial domain (e.g. [2][3][7]), share the disadvantage of a relatively high price and have not yet found market acceptance. The cost-effectiveness is an especially important factor for the compaction support applications, as the asphalt rollers are considered not cost-intensive machines. Still, the correct compaction is critical for the quality of the asphalt pavement.

For this reason, an economical positioning solution based on low-cost DGPS receivers has been studied, and is now implemented in the scope of the OSYRIS project [9]. The intermediate results are presented in this paper.

2. MOTIVATION

CIRC products for the broad market must take the total system cost into account. Compared to the total investment costs of an asphalt roller, additional CIRC equipment is still quite expensive (approximately half of total costs), where most of the price is due to the positioning equipment. Therefore it is justified to consider cheaper positioning solutions.

Required positioning accuracy is the critical factor determining the cost. For surfacing machines, such as asphalt rollers, a 2 dimensional positioning, at a working speed between 5 and 10 km/h is required. For this purpose, Peyret [1]

[1] Institute for Technology and Management in Construction (TMB), University of Karlsruhe, Am Fasanengarten, D-76128 Karlsruhe, Germany, tel. +49 721 608-8222, e-mail: {jurasz | kley}@ tmb. uka.de

recommends decimetre accuracy. Also Do et al. [6] suggest an accuracy better than 15 cm, in order to monitor the overlap between consecutive passes.

Existing DGPS systems can be classified according to the used technology, resulting accuracy and price level as follows:[2]

- Centimetre accuracy Phase DGPS with fixed ambiguities – price range 15-25 k€.
- Decimetre accuracy Phase DGPS with floating ambiguities – price range 3-8 k€
- Meter accuracy Code DGPS – price range 0.1-0.2 k€

As of today, the surveying centimetre accuracy receivers are mostly used for compaction support tasks. Due to the limited market, their price has remained high in recent years. The costly GPS differential phase processing has additional disadvantages: low robustness and long initialisation times. This has to be opposed to the widely applied meter-accuracy DGPS, where the market is well-established, applications are numerous and the price reduction has been achieved using specialised integrated circuits. Moreover, the code measurement is much more robust and is available almost instantly, partly due to the relatively simple processing. This is very important in the context of the road construction site, where the satellite visibility may be limited due to pre-existing bridges, urban canyons etc.

The question arises to what extent the code-differential GPS, possibly aided by inertial positioning and Kalman filtering, may be applied in compaction support context. The experimental tests conducted in 1967 in [7] suggest that extremities of the pavement, the join and edge, tend to receive substantially less compaction. In more recent tests [8], the number of passes at many sections vary between 10 and 50, and the compaction plan has clearly not been kept (see Fig. 1.). Apparently this tendency has not changed during last 35 years.

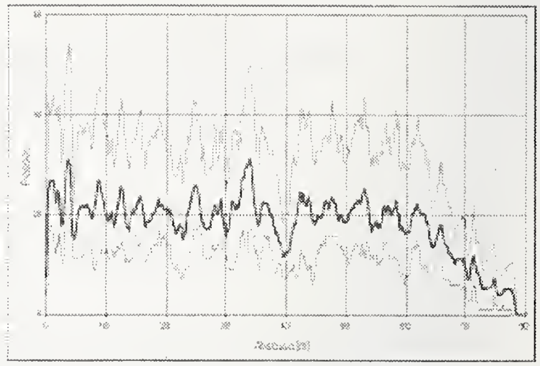


Fig. 1: Minimum, mean and maximum number of drum passes at given curvilinear abscissa measured experimentally on a motorway section.

Clearly, a sub-metre accuracy system would be sufficient to detect and help to correct such major compaction errors. The minimum requirement of the positioning accuracy is posed by the lane definition (typical width 1.8 m). Such level of accuracy can possibly be reached with the Code DGPS. However, it is necessary to investigate the configuration of the sensors and evaluate the resulting accuracy in terms of the man-machine interface (MMI) result.

3. SYSTEM DESCRIPTION

3.1. Configuration

The preliminary study has shown that the application of multiple DGPS receivers on one machine offers several advantages:

- Increased precision can be reached by averaging, as the positioning noise of the receivers is only weakly correlated. The achieved improvement of RMS error was close to $2^{-1/2}$. This suggests that the majority of the error is due to the multipath effects, dependent mostly on the antenna placement.
- The heading of the machine can be determined, also in static conditions.
- As the distance between the antennas remain approximately constant, the reliability of the measurement can be monitored.

For practical reasons the number of the receivers has been limited to two, over the front and rear drum. The additional inertial sensors have been introduced in order to provide independent measurements of the machine speeds: linear and angular. The chosen inertial sensors

[2] As the reference station can be shared, its cost is not taken into account here

are relatively inexpensive and should be integrated in the machine in order to improve the robustness of the system.

Fig. 2. shows the basic structure of the system.

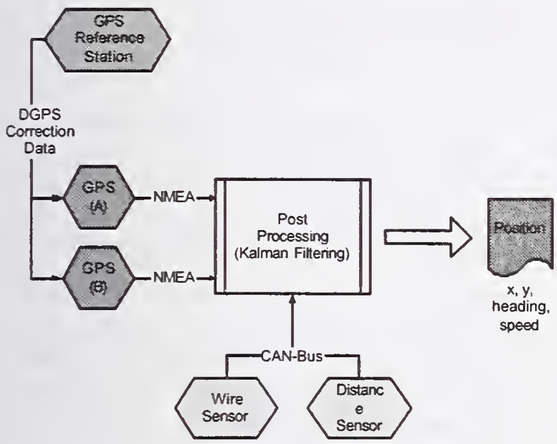


Figure 2: Basic System Structure

The used positioning equipment consist of:

- Two low-cost GPS OEM modules (CEP in DGPS mode < 1 m)[3]
- Two active GPS antennas at the machine extremities
- Wheel odometer
- Optionally: Wire sensor for the articulation angle measurement

The articulation angle measurement can possibly be omitted, as the typical compaction pattern consists of a combination of straight lines, in order to avoid destroying the fresh asphalt surface. The turns are made on already compacted material, with the vibration switched off, and do not have to be counted as passes. However, the articulation angle measurement may be of advantage for curved compaction plans.

The GPS measurements are carried out with 1Hz DGPS mode (with RTCM 2.1 correction messages received from a high quality base station), the inertial sensors work with a measurement cycle of 2.5 Hz.

[3] Manufacturer specifications. CEP, The Circular Error Probability indicates a circle, which encloses 50% of the measured positions.

3.2. Mathematical Modelling

Considering a compactor with an articulation joint, moving in the two dimensional space, the location of the vehicle can represented by the state vector

$$X = [x, y, \phi, v, \omega]^T \qquad (1)$$

where x and y describe the position of the articulation joint, ϕ describes the heading (yaw), v and ω denote respectively the translational and rotational speed (see Fig. 3.). Both drums can be modelled by a stick of a length w (equal to the drum width) attached at to the joint. This simplified model can still accurately represent the longitudinal overlaps.

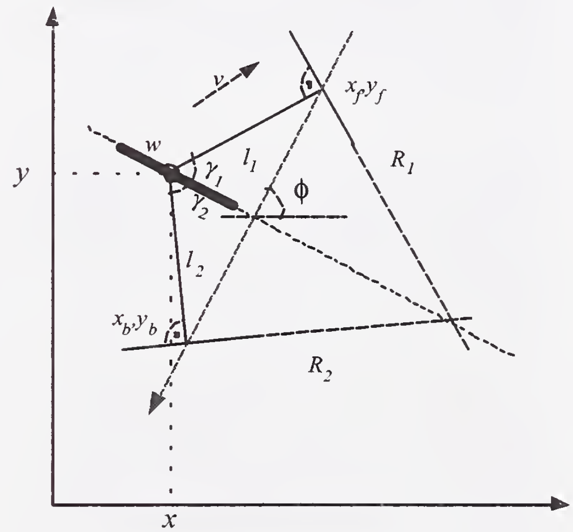


Figure 3: Model geometry

Assuming constant speeds and no lateral slip, and taking into account the articulation angle, the kinematic model is represented by the nonlinear model

$$\dot{X} = F(X) = \begin{cases} \dot{x} = v\cos\phi \\ \dot{y} = v\sin\phi \\ \dot{\phi} = \omega \\ \dot{\omega} = v(\gamma - \pi) \text{ or } 0 \\ \dot{v} = 0 \end{cases} \qquad (2)$$

The measurement vector Z includes the position measurement from the front and the rear GPS, placed l_1 and l_2 from the articulation joint, above the front and the rear drum of the roller.

Moreover the angle and speed measurement are used.

$$Z = \left[x_f, y_f, y_b, y_b, \phi, v, \gamma\right]^T \tag{3}$$

The discrete system equations can be then written as

$$\begin{aligned} X_k &= F(X_{k-1}) + V_k \\ Z_k &= H(X_k) + W_k \end{aligned} \tag{4}$$

where and F and H are system and measurement functions and V, W are additive system and measurement noises.

Based on the estimated speed, a Boolean estimate of the static/kinematic mode can be produced. The passes of the machine are counted only in the kinematic mode.

3.3. The Unscented Kalman Filter

The observations are processed with an Unscented Kalman Filter (UKF). The UKF for nonlinear estimation [4] is based on the Unscented Transformation, presented by Julier and Uhlmann [5].
Consider a recursive estimation

$$\hat{X}_k = X_{k,pred} + K_k \cdot \left(Z_k - Z_{k,pred}\right) \tag{5}$$

for the optimal minimum mean-squared error estimate for X_k, assuming that the prior estimate $\hat{X}_k$ and the state observation Z_k are Gaussian Random Variables.
The UKF extends (5), redefining the state random variable as the concatenation of the original state and noise variables [3]. An unscented transformation sigma point selection scheme is applied to the new augmented state random variable to calculate the corresponding sigma matrix [5]. The UKF algorithm can be found in [4].

Compared to the traditionally applied Extended Kalman Filter (EKF), the UKF offers better performance and adapts better to non-gaussian noise present in this non-linear system. Moreover, as the explicit linearisation is not required, it is easy to test alternative models and filtering structure. The disadvantage is increased computation cost in terms of CPU time and memory.

The full Kalman step is calculated in a 1 Hz cycle, with each new GPS measurement. Assuming a working speed of 2 m/s this results in a real-time uncertainty of 2 m on the MMI, additionally influenced by the GPS latency (< 0.5 s corresponding to 1 m). For these reasons a filter with a 2-5 Hz prediction step and a 1 Hz correction step has been considered.

4. EXPERIMENTAL VALIDATION

To conclude on the accuracy of the presented positioning solution, several static and dynamic tests were performed and evaluated.

As no asphalt roller was available for the first test period, a wheel loader with similar kinematic characteristics was chosen (Fig. 4.). The two GPS antennas are mounted at the front shovel and on a framework at the back, where also the wheel-distance sensor is fixed.

To reduce the multi-path errors, a metallic disc, diameter ~40 cm, is mounted under both antennas. The last tests were performed, using an additional framework at the back, to reach an antenna height of 2.50 m. Also the shovel at the front was lifted, to reach the same antenna height as at the front.

Figure 4: The machine and equipment used in the preliminary tests.

The raw data from the GPS and the inertial measurements are processed, using an UKF filter, with the kinematic model, presented above.

The test results (Fig. 5.) showed that in the static conditions one can observe a drift of the measured position. To account for this effect UKF switches to a static model, represented by:

$$\dot{X}_{static} = F(X) = 0 \tag{6}$$

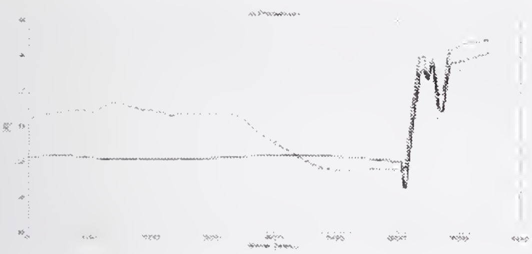


Figure 5: X-coordinate of the position: raw front measurement, raw rear measurement, processed by the Kalman filter

In order to establish the reference for the obtained accuracy, the recorded raw GPS observables (code pseudoranges and integrated carrier phases) have been processed using commercial GPS post-processing software. Its performance can be judged similar to the best GPS receivers available. In this way centimetre-accurate trajectories of both antennas have been measured.

The obtained trajectories together with the heading information have been converted into ribbons [10] and presented as pass maps in the Fig. 6. Except of the overlaps, both maps are qualitatively very similar and give the operator the required overview of the performed compaction work. The major omissions and the tendency to over-compaction in the middle of the pavement and under-compaction close to the edges can be easily detected.

a.  b. 

Figure 6: Pass maps observed in the experiment. a. Code - measured on-line based on code-differential processing, b. Phase – post-processed using phase observables

The comparison between the Code and the Phase results has brought the following results:

- The Code-DGPS modules estimate driving/stopped state and perform position, velocity, time (PVT) type filtering on their own. This has an effect of additionally smoothing the result. However, the movements at low speed (< 2 km/h) are filtered out. This effect can be observed on the Code pass map as an early stop condition, leading to an underestimation of the compacted surface. This effect is not critical in the practice, as the breaking/acceleration occurs with the vibration switched off and should not be counted as a valid pass.
- At the moment the absolute accuracy of the 2 Code DGPS system can be judged at about 80-90 cm, after removing an observed constant offset between the Code and Phase results. This is sufficient to obtain clear lanes in the pass map. However, longer tests are required to confirm this finding.
- The GPS signal reflections caused by the machine and mounting seem to have great impact on the multipath error and the final result. The validation on the target machine is necessary
- The performance of the filter with and without the articulation angle measurement was similar. For studied trajectories it is possible to remove it from the filter structure. However, due to varying width, this measurement is required for the machines with 2 articulation joints.

5. CONCLUSIONS

The presented positioning solution reaches good performance in static and dynamic mode. The pass maps obtained with the low-cost system are qualitatively very similar and allow for detection and correction of the major compaction errors identified by the experimental tests. Long term tests on the target machine are necessary in order to tune the filtering algorithms and fully asses the accuracy and robustness of the proposed solution.

The availability of suitable cost-effective and robust DGPS positioning solutions should improve the acceptance of compaction support systems in the near future.

The full scale tests of the presented algorithm are scheduled in Summer and Autumn 2002.

Acknowledgements

The work described in the paper was carried out in the framework of the European Project OSYRIS: Open System for Road Information Support (Growth No. GRD1-1999-10815), 2000-2003 [9].

6. REFERENCES

[1] F. Peyret. *The paradigm of Computer Integrated Construction applied to Civil-Engineering*, Proc. 16th ISARC, Madrid 1999.

[2] F. Peyret, J. Jurasz, A. Carrel, E. Zekri, B. Gorham: *The Computer Integrated Road Construction project*, Automation in Construction 9 (5-6) (2000) pp. 447-461.

[3] D. Bouvet, M. Froumentin, G. Garcia. *A real-time localization system for compactors*, Automation in Construction, 10 (4) (2001), pp. 417-428.

[4] E.A. Wan, R. van der Merwe. *The Unscented Kalman Filter for Nonlinear Estimation*, Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC), IEEE, Lake Louise, Alberta CA, 2000.

[5] S.J. Julier, K. Uhlman. *A new Extension of the Kalman Filter to Nonlinear Systems*. Proc. Of AeroSense, 11th Int. Symp. On Aerospace/ Defence Sensing, Simulation and Controls, 1997.

[6] W.-S. Do, A.A. Oloufa, H.R. Thomas. *Evaluation of Global Positioning System Devices for a Quality Control System for Compaction Operations*. Transportation Research Record 1675, Paper No. 99-0139, pp. 67-74.

[7] M.J. Kilpatrick, R-G. McQuate: *Bituminous pavement construction*, Federal Highway Administration, Washington, DC., 1967.

[8] Geodynamik Internet site, online resource at http://www.geodynamik.com.

[9] Osyris Internet site, online resource at http://www.osyris.org.

[10] J. Jurasz. *Universal digital environment model for computer integrated road construction*. Proc. 18th ISARC, Cracow 2001.

ON-BOARD DATA MANAGEMENT STRUCTURE FOR ADVANCED CONSTRUCTION MACHINE SUPPORT

by

Jarosław Jurasz[1], Agata Ligier[1], Alfons Horn[2], Jochen Wendebaum[1]

ABSTRACT: A flexible on-board data management structure for a construction machine and interfaces between its elements are discussed. The solution, which grants the required flexibility and enables advanced, distributed functions is discussed on the example of the road paving process.

KEYWORDS: CANopen; digital work documentation; computer integrated road construction; road product model; site information systems; OPC; XML

1. INTRODUCTION

Nowadays the construction machines are increasingly often equipped with the on-board computers, which support the operator and take over the control and documentation functions. Their application is especially promising in civil engineering, where the machines perform repeatable and well-defined tasks, e.g. laying, compacting, grading. Thanks to advances in positioning technology, on-board communication and following the need of the users, who wish to actively participate in the quality control (Build-Own-Operate etc.), the on-board IT plays increasing role on the modern construction site.

Major part of the requirements set for the on-board IT is due to the need for applicability to different machines and worksites. This includes varying configuration of positioning equipment, sensors and measurement systems coming from different providers.

With the increasing number and extending functionality of CIRC systems in existence, the problems of interoperability and standardisation start to play a major role. Those topics are addressed in the Osyris project, in the frame of which this work is carried out.

The goal set by the authors is to design a flexible on-board data management structure for a construction machine, equipped with standardised interfaces. The article is structured as follows: after discussing the context and the requirements for the on-board structure the contents of the information is investigated. Then the Osyris layered on-board structure is presented, followed by two examples of advanced, distributed functionality.

2. CONTEXT

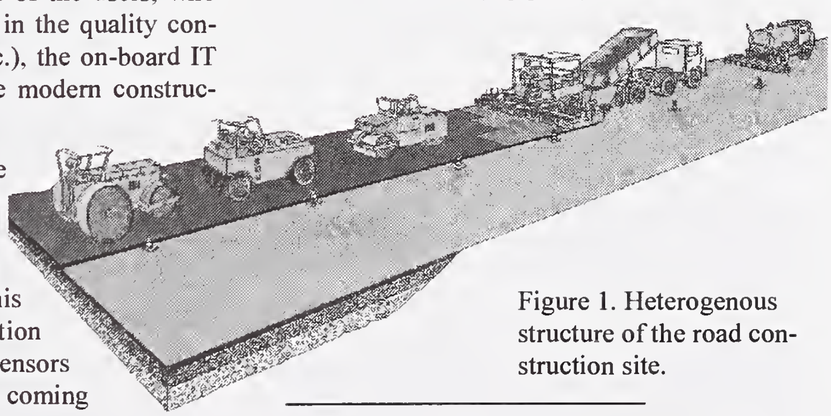


Figure 1. Heterogenous structure of the road construction site.

[1] Institute for Technology and Management in Construction (TMB), University of Karlsruhe, Am Fasanengarten, D-76128 Karlsruhe, Germany, tel. +49 721 608-3884, e-mail: {jurasz | ligier | wendebaum}@tmb.uka.de

[2] Moba Mobile Automation GmbH, Vor den Eichen 4, D-65604 Elz, Germany, tel. +49 6431 9577-132, e-mail: horn@moba.de

The major challenge for the Osyris on-board data management structure is due to the heterogeneous information structure of the worksite (see Fig. 1.), featuring:

- Separate processes (e.g. earthmoving, laying, re-profiling, transport), often performed by separate organisations,
- Few standardised interfaces: GPS NMEA standard[3] is one notable exception,
- Differently equipped machines coming from different manufacturers,
- Varying configuration (ad-hoc task allocation and team definition, add-on sensors, different operating modes).

Clearly a separation is needed between the managed process information and the details of the algorithm used to obtain or process it. To assure flexibility and extensibility of the implementation we propose to fix only data storage structure and its interfaces. In this way the information can be accessed and/or provided from any component without revealing the details of how the information was acquired. As the machines work in team, wireless real-time interfaces to the other machines have to be taken into account.

Nowadays 80..90% of the projects conducted by European road contractors are various maintenance tasks. The typical limited maintenance configuration (small paver and 1-2 small rollers) requires simple and cost-effective solutions.

3. INFORMATION CONTENTS

The contents of the managed information is defined in the Osyris Functional Design. A distinction between volatile real-time and static information can be made here. The most important data categories managed on-board are (see Fig. 2.):

- **Design** – a description of a target road in a suitable digital form (what). Remains static during the project execution.
- **Mission** – a description of the task to be performed by a given machine (how). Remains static for the given mission, changing typically daily.
- **Machine state** consists of position, tool geometry and the process data, which vary in real-time. It can be used to derive the achieved work.
- **Target machine state** is a subset of machine state subject to automatic control, derived from design, mission and current machine state, for example designed elevation at the given point.
- **Achieved work** is a record of work execution. It is gathered in real time and remains static after the work execution.

All the data categories listed above can be expressed as parameter values assigned to a geometry, where design, tool or achieved geometry can be used to carry the parameters. Depending on the geometry they are assigned to, the same parameters can be used to specify the actual or target values. The managed parameters can be outlined as follows:

- Position: Geographic, Curvilinear coordinates, linear and angular speed
- Inherent properties of the material: temperature and contents
- Geometrical properties: thickness, evenness, volume, level deviation
- Process and machine parameters: vibration, tampering, hydraulic pressure.
- Ambient conditions (temperature, wind, sunlight)

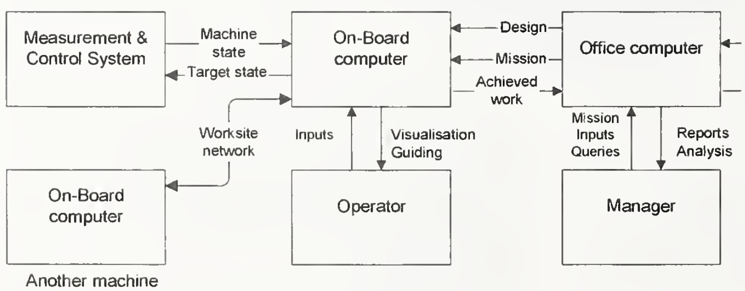


Figure 2. Functional decomposition of an OSYRIS system

Clearly it is not possible to define an exhaustive list of the parameters, so the definition of the additional parameters must be possible at the runtime. The identification of parameters can be performed using standardised names or codes. In any case it is very important to clearly define the meaning and units of the parameter values.

[3] Although many GPS receivers offer best performance using custom interfaces.

The piecewise-linear ribbon structure[4] [1] can be used to efficiently represent all data categories listed above. In the case of achieved work description, the tool geometry with assigned parameters describing state can be used as a ribbon generator. The ribbons introduce the natural ordering of the target parameter values along the road axis, so that extrapolation principle can be used to limit the number of managed parameter values.

It is important to note that multiple values of one parameter may coexist at given time. For example the following values of machine speed need to be managed concurrently on-board:

- measured by absolute positioning device, e.g. GPS or Robotic Total Station
- measured by local positioning device, e.g. encoder
- elaborated by data fusion algorithm, e.g. Kalman filter
- specified as target for the mission.

4. LAYERED ON-BOARD STRUCTURE

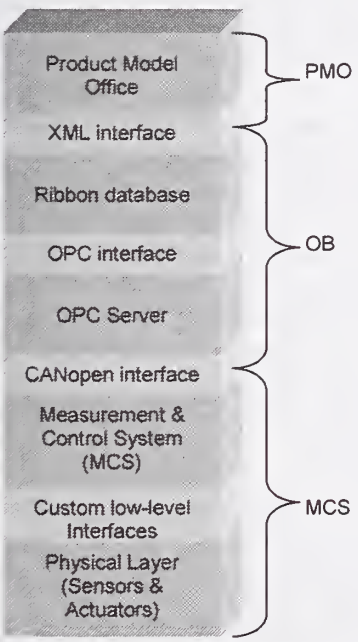


Figure 3. Osyris layered on-board structure

The Osyris system is based on components. In comparison with monolithic systems (e.g. CIRC [3]) this approach enables easy reconfiguration and provides framework for specialised components which provide advanced services (e.g. soft sensors).

The on-board system layers are presented in the Fig. 3. The abstraction level, amount of managed data, and required performance depend on the specialisation level:

PMO – responsible for permanent storage, mission preparation and work analyses [2].

OB – responsible for visualisation and storage functions, best effort (not guaranteed) real time behaviour.

MCS – responsible for guaranteed real time interfacing, filtering and control functions. Multiple MCS devices are possible, although as of today an architecture with single, central MCS is the most economical solution.

4.1. CANopen interface

The Osyris CANopen interface approach is based on concept of a device profile, describing the measurement and control objects provided by the real time measurement and control systems. The fast exchange of measured data between the MCS and the on-board computer is guaranteed by the efficient CANopen communication layer (so called communication profile). The CANopen communication profile is already standardised.

Former approaches to the interfacing to the measurement and control systems on construction machines were based on lower level solutions (e.g. CAN) and, as the interface description was hardcoded, changes in the system structure required a lot of modifications on software and hardware level, which led to additional costs and less reliable systems.

Therefore the Osyris solution promotes a standardised communication for construction machines. As a part of the proposed solution the standard connector is also presented. The advantages of standardised devices are numerous. It encourages the manufacturers to produce standard measurement and control devices with their own technology hidden behind the standard interface.

Furthermore the open solution allows contractors to use construction machines coming from different manufactures and connect them

[4] The ribbon is created by moving a planar figure called generator along a curve called axis [1]

to one on-site management system, without any adaptations or modification in the MCS.

The central part of the device profile is the object dictionary description (an extract is shown in Table 1). The object dictionary is essentially a grouping of objects accessible via CANopen in an ordered, predefined fashion. Each object within the dictionary is addressed using a 16-bit index and 8-bit sub-index. The Object Dictionary contains only few mandatory items. They can be accessed according to the worksite requirements and a current configuration. The MCS not only offers but also requires items. The availability of items is resolved by the on-board computer in runtime.

Table 1. Extract of the paver device profile

Index	Object	Description
6000	Type of machine	Kind and type of machine
6010	MCS functionality	Indices of items supported by particular implementation
6020	Event	General (start, stop etc) and machine-specific (vibration on/off etc) events
6100	Position	The geographic position of machine's tool in local coordinate system (E,N,H)
6101	Angle Position	The attitude of the tool
6102	Curvilinear Co-ordinates	The coordinates of the tool in the local curvilinear system
6103	Level Deviation	The levelling error
6200	Thickness	The thickness of the laid layer
6300	Screed width	The width of the screed and its extensions
6310	Volume	The volume of the laid material
6500	Material core temperature	The temperature of the laid material
8010	MCS wish list	Indices of items required by the MSC

4.2. OPC layer

The standardised interface to the actual machine state parameters at the level of on-board computer is defined with help of the OPC.

OPC (OLE for Process Control) is a standardised set of interfaces, based on OLE/COM and DCOM technology, for open software application interoperability between control applications, field systems and devices, and business/office applications [4]. Osyris uses a subset of OPC called Data Access. The data exchange is based on the server (provider)-client (consumer) principle (see Fig. 4.).

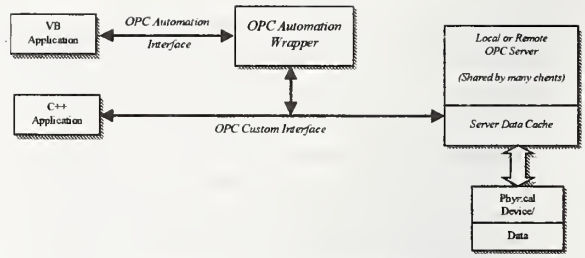


Figure 4. OPC Architecture

OPC provides a snapshot of current machine state as a hierarchy of named items together with the time stamp and quality information. The items have standardised names and are mostly numeric values (vector and string are also possible). The scaling between metric (SI) units and device units is performed by the Osyris OPC Server.

Osyris OPC namespace is divided in the following groups, mostly according to the origin:

- CanOpen – machine state as received from the MCS
- CanOpenOut – volatile target machine state (e.g. level deviation) elaborated by the on-board computer, transmitted to the MCS
- Designed – target state (e.g. thickness) as defined in the mission
- Pos – machine position elaborated by the positioning component
- Manual – parameter values may be overridden manually by the operator
- Default – for many parameters, e.g. Temperature or Speed sensible default values can be defined here, and used in case of lack of sensors
- Process-specific groups, e.g. Material, CES (Compaction Expert System), CM (Cooling Model)

In the Osyris implementation there is only one Osyris OPC server, which allows reading and writing clients (no arbitration for writing is guaranteed at the OPC level). The Osyris implementation uses the OPC concept not only for client-device communication, but additionally (which is an extension to the OPC standard) as a storage for the current machine state. Also clients which perform parameter estimation (soft sensors) can write into current state. Writing to the device (in this case MCS) is allowed only in the CanOpenOut group.

In addition to the standard OPC, a concept of the 'unqualified items' has been introduced. As already mentioned, there can exist more than one value for a parameter. If the client component is interested in the 'best' value, it subscribes to the unqualified item (without specifying the group) and the server decides which one it is, based on the priority list and current qualities. (see Fig. 5.)

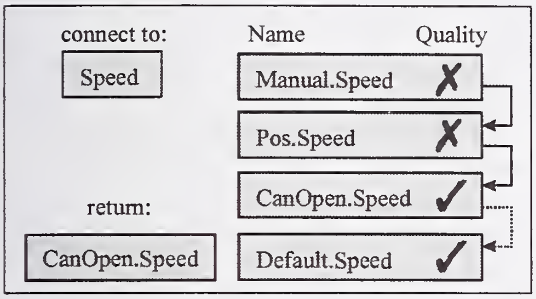


Figure 5. Example of Osyris OPC connection using an unqualified item name.

Further extension of the OPC standard is the meta-information contained in the 'Events' OPC group. This concept is used to notify the client that the best source of data has changed. In this way a fallback mechanism is implemented.

The OPC standard includes also a time-referenced interface to the past data, so called historical OPC. Unfortunately it cannot be employed in the CIRC context, as space reference is missing. OPC can be also accessed remotely via DCOM. It is not used directly by the Osyris framework, but can be advantageous for debugging, scripting or Web clients.

4.3. Ribbon database

As described in [1], ribbons (Fig. 6.) can be used as a permanent storage for design, mission and achieved work.

For each machine the tool geometry (defined in mission or measured) is a machine ribbon generator. There exist one dynamic ribbon for each machine, containing data gathered in real time (for more complex tool geometries multiple ribbons are foreseen). The recorded position is represented by a ribbon diagonal, with the parameters assigned to it.

Ribbons provide following services:

- Map visualisation of the parameters
- Storage for cooling model and other algorithms
- Geometrical searching and iteration
- Parameter interpolation, e.g. elevation, thickness interpolation
- Curvilinear transformation.

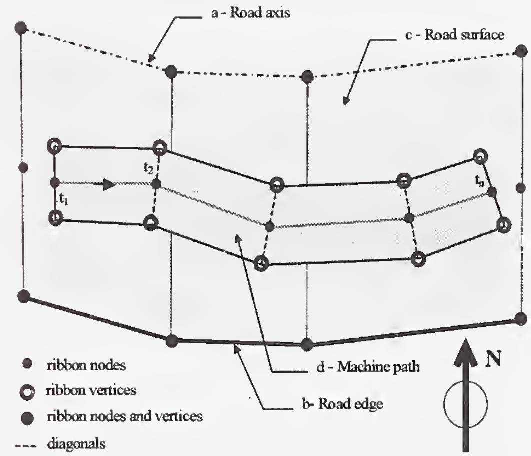


Figure 6. Sample ribbons in geographic coordinates representing: a - road axis, b – road edge, c – road surface and d - machine path

4.4. XML interfaces

The interfaces between the on-board and the Office components are based on the XML standard. XML standard is wide spread in Internet applications and many tools are available. A concept central to XML is the separation of the information contents and format. The information contents of the geometry-based Osyris XML files is based on ribbon concept.

At the beginning of work the on-board must be supplied with mission information (worksite geometry, designed values of the work parameters like speed, fleet configuration). At the end of the work the achieved information has to be exported. There is also a request-response schema, allowing the transfer of the current work state at any time. The detailed schema has been defined for each exchange file format.

5. DISTRIBUTED OPERATION

5.1. Wireless communication

Wireless communication between machines is based on a reliable multicast implemented upon Internet datagram (UDP) protocol. Typically WaveLANs are used as physical transport. Ribbons with parameters are automatically synchronised among the machines. This property can be used to implement distributed algorithms (see the following subchapters).

5.2. Distributed pass counting

One of the most valuable information concerning the quality of compaction comes from the pass counting. Given the past position of the compactor, the number of the done passes can be calculated for each point of the surface, resulting e.g. in a compaction coverage map. On the most worksites at least two compactors are working together. The coverage map is only then useful if it contains passes from all the machines, and this is guaranteed by the wireless communication algorithm.

5.3. Cooling model

Another important parameter to take into account during the asphalt laying is the temperature of the asphalt layer. Attempts to measure the surface temperature of the asphalt at the compactor using infrared sensors are unreliable due to the high influence of the wind speed on the surface temperature.

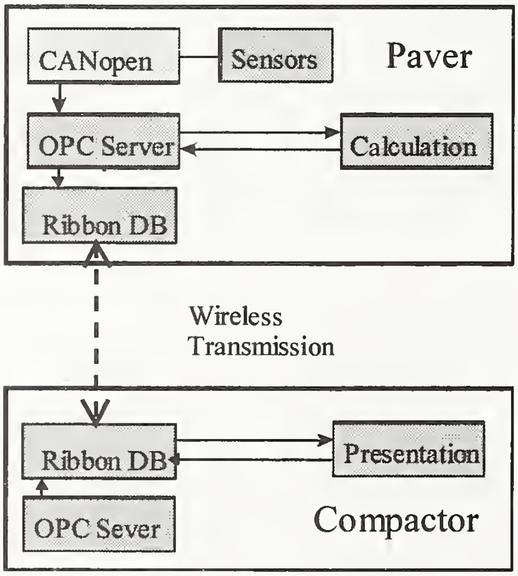


Figure 7. Cooling model as an example of a distributed application in the OSYRIS system

To calculate the evolution of the temperature inside an asphalt layer, a mathematical model has been used. The model parameters like *layer thickness* and *asphalt temperature at laying* can only be measured at the time of paving on the paver, but the results of the core temperature calculation are most interesting for the compactor operator, giving him the very valuable information about the time left to finish the compaction (Fig. 7.).

6. CONCLUSIONS

The presented on-board structure is universal: all the relevant process parameters are represented in uniform and coherent way and can be transparently accessed on all the levels of the system. Pre-existing standards have been used as a base for the Osyris interfaces.

The presented solution grants the required flexibility and enables advanced, distributed functions, for example distributed pass counting and real-time simulation of the asphalt cooling (cooling model).

Validation of the proposed framework on the worksite is planned for Autumn 2002. The specifications of the interfaces will be published at the end of the project.

Acknowledgements

The work described in the paper was carried out in the framework of the European Project OSYRIS: Open System for Road Information Support (Growth No. GRD1-1999-10815), 2000-2003 [5].

7. REFERENCES

[1] J. Jurasz, Universal digital environment model for computer integrated road construction. Proc. 18th ISARC 31-36, Cracow 2001

[2] A. Ligier, J. Fliedner, J. Kajanen, F. Peyret, *Open System For Road Information Support*. Proc. 18th ISARC, Cracow 2001

[3] F. Peyret, J. Jurasz, A. Carrel, E. Zekri, B. Gorham: *The Computer Integrated Road Construction project*, Automation in Construction 9 (5-6) (2000) pp. 447-461.

[4] OPC Foundation at http://www.opcfoundation.org

[5] Osyris Website at http:// www.osyris.org

OFFLINE PATH PLANNING OF COOPERATIVE MANIPULATORS USING CO-EVOLUTIONARY GENETIC ALGORITHM

by

M.S.Ajmal Deen Ali[1], N. Ramesh Babu[1] and Koshy Varghese[2]

[1]*Dept. of Mechanical Engg., I.I.T. Madras, India, email: nrbabu@iitm.ac.in*
[2]*Dept. of Civil Engg., I.I.T Madras,India, email: Koshy.Varghese@asu.edu*
(Currently on Sabbatical as Visiting Eminent Scholar at Del E. Webb School of Construction Arizona State University, Tempe, AZ)

ABSTRACT: This paper presents a new path planning approach using Coevolutionary genetic algorithm (CGA) for automating the path planning of two Cooperative construction manipulators. A methodology based on the concept of configuration space (C-space) technique in conjunction with the Coevolutionary genetic search is used for generating the path. The paper proposes a method for finding the minimum distance and collision free path using CGA. It uses an interference detection algorithm that runs in parallel with CGA to check collision between obstacle and cooperative cranes during path planning. The effectiveness of CGA is compared with other search approaches like A* and Genetic Algorithms GA.

KEYWORDS: Cooperative manipulators, Path planning, Coevolutionary genetic algorithm

1. INTRODUCTION

Cooperative manipulators are being widely employed in construction and assembly for handling medium to heavy objects. Path planning of cooperative manipulators is quite different from path planning of a single manipulator. When the object to be lifted is too heavy or large for a single manipulator, options such as specially assembled equipment such as jacking systems and cooperative use of multiple medium capacity manipulators can be a suitable alternative.

Path planning techniques are classified based on the criteria such as dimension of space, mobility of manipulator and obstacles, representation of space and nature of information gathering [1-6]. Trajectory planning of two manipulators, cooperating to manipulate the same object, was solved using genetic algorithm (GA) [7]. Evolutionary algorithm has been used to find time-optimal trajectories of two coordinating manipulators sharing the same work space [8]. Customized genetic operators such as analogous crossover, which suits certain problem situations, was also designed [9]. Recent efforts include the development of natural systems that can capture key features such as Co-evolution and Life Time Fitness Evaluation, for solving path-planning issues [10].

Earlier attempts for path planning include Hill climbing search, A* search and Genetic Algorithm (GA) search [11, 12]. These approaches have certain disadvantages like excessive computational time, memory requirement and non-optimal paths. Recent works utilize the concepts of co-evolutionary genetic algorithm (CGA). Two populations constantly interact and co evolve in case of CGA.

A 2-D path-planning problem using CGA was attempted [10]. This paper presents a new approach to solve the 3-D path-planning problem of cooperative manipulators using CGA.

2. PROBLEM STATEMENT

This section presents the details of a path-planning problem involving two spatial cooperative manipulators, each of 3 DOF with hinged base. Two important considerations of cooperative manipulators path planning are:

(a) *Ensuring cooperation between the two manipulators during lifting.*

(b) *Handling the computational complexity of the problem based on the DOF of the cooperative manipulator system and its movement in obstacle clustered environment.*

a) Ensuring cooperation between the two manipulators during lifting

Cooperation between the two manipulators is ensured by (i) Nature of movement of cooperative manipulators. (ii) Spatial distance between the hook ends of the manipulators during cooperative lift. (iii) Altitude difference between the hook ends of the manipulators and (iv) Hoist limit evaluation for both manipulators.

Nature of movement represents the movement between the two arms, which can be either synchronous or asynchronous. A synchronous movement refers to the identical movement of different joints of cooperative cranes between two steps. In case of asynchronous movement, the movement between the two successive steps will not be the same.

The spatial distance between the boom tips of the manipulators is kept within object length ± 2 units. When the spatial distance between the boom tips of both manipulators is different from the object length, the sloping of load line occurs. This can be computed by (Figure 1).

$$\beta = \tan^{-1}\left[\frac{D}{HL}\right] \qquad (1)$$

where 'HL' - hoisting length of the hook and 'D' - Off-lead i.e. displacement of the hook from boom tip. Let 'W' be the vertical component of the load transferred from the object to the manipulator and 'W_a' the actual load acting along the load line of the crane. 'W_a' is calculated by

$$W_a = W / \cos\beta \qquad (2)$$

$(W - W_a)$ is the additional load transferred to the crane due to the slope 'β' of the load line. This results in additional load transformation to the manipulators.

If the altitude difference (H1-H2), as shown in figure 2, exists, different loads will be acting on the hook ends. This results in more payload acting on a particular manipulator. In order to limit the extra load on the crane, an altitude difference of '3' units is considered.

Hoist limit evaluation represents the safety to be considered in hoisting the rope in either direction i.e. up or down from the ground level. Minimum limit is considered as '0' unit i.e. at ground level and maximum position is boom tip position, considered above ground level, which keeps on varying depending on the luffing angle.

b) Handling the computational complexity of the problem based on the DOF of the cooperative manipulator system and its movement in obstacle clustered environment

Collision of the cooperative manipulator system in the obstacle-clustered environment can occur due to (i) manipulator-1 colliding with obstacles, (ii) manipulator-2 colliding with obstacles, (iii) object colliding with obstacles and (iv) manipulator-1 colliding with manipulator-2.

The feasible movement of the manipulators in the obstacle-clustered environment is ensured by means of interference detection algorithm that assesses the different configurations of the manipulator for collision. The collision checking is done by means of interference checking algorithm that performs interference check between line and plane segments. For this purpose, the obstacle is represented by many plane segments and cooperative manipulator system is represented by many line segments.

A test problem considers two spatial cooperative manipulators performing asynchronous movement. Both manipulators are identical in size and shape. Table 1 shows the configuration of different arms. The upper and lower bound movement for different arms are shown in Table 2.

3. SEARCH METHODOLOGY

This section covers the details of modeling and implementation for 2x3 cooperative manipulator system using A*, GA and CGA.

3.1 A search in open C-space*

A* search is a free search in the open C-space with on-line feasibility check [12]. The main advantage of this search is that it has the ability to go back to any node which was visited earlier. The step size values of swinging, luffing and hoisting for generating neighbors are 5, 5 and 1.

3.2 GA search in open C-space

GA search is a free search in the open C-space with on-line feasibility check [11]. The population consists of 250 strings. Each string in GA represents movements of the manipulator from pick to place location. A typical string with fifteen intermediate configurations between pick and place location is shown in figure 3.

3.3 CGA

In this approach, two populations known as solution population and test population continuously interact with each other to evolve an optimal solution.

a) Solution population: Path representation as a string

The solution population consists of 200 strings. Each string represents the movement of manipulator from pick to place location (Figure 3). The number of steps between pick and place location is taken as fifteen.

b) Test Population

The test population consists of test conditions known a priori (Figure 1 and Figure 2). It consists of (1) spatial distance between the hook ends of the manipulators i.e. object length ± 2 i.e. 15+2 and 15-2, (2) altitude difference between the hook ends of the manipulator i.e. (H1-H2) = '3' units, (3) hoist limit evaluation varies according to boom luffing angle and (4) collision of the cooperative manipulator system with the environment that consists of (i) manipulator-1 colliding with obstacles, (ii) manipulator-2 colliding with obstacles, (iii) object colliding with obstacles and (iv) manipulator-1 colliding with manipulator-2.

c) Encountering

In CGA encountering is the process in which one string from solution population and randomly chosen test conditions from test population interact with each other to produce a better offspring. It consists of three stages.

STAGE 1

Fitness is found for all the individuals in both the populations by subjecting them to encounter '3' randomly selected individuals from the other population. A solution receives a payoff of one if it satisfies the test. Otherwise, it receives a zero. The opposite is true for test. It gets a payoff of one if the solution encountered does not satisfy a test. Each individual - test or solution - has a history, which stores the payoff resulting from such an encounter. The fitness of an individual in solution population is estimated by

$$F_s = P\ (x)\ [1+(1/H_s)] \qquad (3)$$

where P (x) is the objective function and H_s is the total payoff for that individual. The objective function is defined as the sum of square of absolute differences of identical joint angles between successive configurations for all the joints of the manipulators as the cooperative manipulator system moves from pick to place location. It is estimated by

n-1 m

$$P(x) = \sum_{i=1} \sum_{j=1} |\theta_{i+1,j} - \theta_{i,j}|^2 \qquad (4)$$

where 'n' represents the number of configuration sets and 'm' represents the number of joint parameters required to define a unique position of the cooperative manipulator system and '$\theta_{i,j}$' is the value of joint angle of j^{th} link in i^{th} configuration set. The fitness of an individual in the test population is estimated by

$$F = H_t. \qquad (5)$$

where H_t is the total payoff for that individual. An individual according to their fitness value i.e. minimum distance and payoff for first population and only payoff for second population, is arranged in descending order in their respective population

STAGE 2

In this stage, one fittest string from the first population and three randomly selected test conditions from the test population are subjected to encounter. The selection of this string and test conditions are biased towards highly ranked individuals i.e. the fittest individuals are more likely to be selected. The result of such an encounter is '1' if any test is satisfied (or) '0' in case of violation. The fitness (or) maximum pay-off for this string is calculated. According to its fitness value, it is ranked in descending order in the first population. A similar operation is performed on the test population also. Since both the populations are sorted based on their fitness values, an individual might move up and down in its population as a result of the update of its fitness.

STAGE 3

In this stage, the conventional GA process is followed. Two strings from the first population are selected based on their fitness. An offspring is created by the process of adaptive crossover. Adaptive crossover is illustrated in figure 4, i.e. the same intermediate configurations on both parent 1 and parent 2 will be checked for their payoff value. The particular configurations, with highest payoff, will be inserted in to the same configuration in the offspring. This cycle is repeated for the remaining configurations until all the configurations in the offspring are filled.

Mutation is applied in an adaptive manner with a probability of '0.1'. Adaptive mutation is implemented in order to reduce the angular displacement between adjacent configurations as well as to bring the cooperation between manipulator 1 and manipulator 2. All the joint angle positions are subjected to this probability. If they are satisfied, they will undergo adaptive mutation as shown in Figure 5. For example, the swing position as marked by arrow '1" is to be subjected to adaptive mutation (assumed to be the j^{th} position) then the j-6^{th} (arrow 2) and $j+3^{rd}$ position (arrow 3), both are swing positions, will be considered. A random number will be generated between the swing positions considered. Similarly if luffing position is considered as j^{th} position, then the j-6^{th} and $j+3^{rd}$ positions i.e. luff position, will be considered for mutation. A random number will be generated between these points. The value generated is inserted in the j^{th} position. This process is applied to all the joint angle positions, which satisfy the mutation probability condition.

Figure 6 shows the methodology adopted for finding a feasible collision free string with CGA. The fitness of offspring is estimated as the sum of payoff received from encounter with three-selected test conditions. The offspring is then inserted into the solution population based on their fitness values. To accommodate this new offspring, the lower fitness value string in the solution population is deleted. The procedure adopted in STAGE 2 and STAGE 3 is continued until the fitness of string in the solution population remains almost the same in ten consecutive generations.

4. RESULTS AND DISCUSSIONS

The manipulators shown in the section 2 are considered for cooperative path planning with the proposed CGA approach. In order to asses the effectiveness of this approach, two different approaches such as A* and GA for path planning were also considered and these approaches were developed using C++

programming language and implemented on the same platform, i.e. 333 MHz Pentium II processor PC with 128 MB RAM with Windows NT operating system. The computation time for finding the feasible path from pick to place location using different approaches like A*, GA and CGA, depends on the lift path and position of the pick and place location in the obstacle clustered environment i.e. lifting an object vertically up may be simpler to compute than another path. For path planning of cooperative manipulators, C-space approach was adopted for representing the position of the manipulators. Performance of A*, GA and CGA for path planning of 2x3 cooperative manipulators is assessed and the results are presented in Table 3.

*4.1. A*search*

The minimum incremental movements of different joints of manipulator are shown in Table 4. A* search could determine the feasible path from pick to place location in 30 intermediate steps. The time taken for finding the feasible path is 320 minutes. The minimum distance, in terms of linear movements, from pick to place location is 73 units. If A* is allowed to search the C-space exhaustively i.e. with 1 degree increment, the time taken will be more than 2000 minutes. The disadvantage with this step angle increment is that pick and place angle has to lie within the increment angle of search; otherwise the search will never give a feasible solution.

4.2. GA search

Total number of joint angles to represent a unique configuration is 6; Range for different joint angles swing: 0-360 degrees: luffing: 30-80 degrees: hoist: 0-39 units; Number of configuration sets representing a string excluding pick and place location: 15.

For generating the path with GA, the number of intermediate configurations between pick and place location is fixed as fifteen. The time for computation of feasible path with GA is 183 minutes. The minimum distance in terms of linear movements from pick to place location is 79 units. GA computes the collision for all the configurations in the population, due to which a considerable time was spent.

4.3. CGA search

Total number of joint angles to represent a unique configuration: 6; the arm configurations and its lower and upper bound values are shown in Table 3 and Table 4. Number of configuration sets representing a string excluding pick and place location: 15.

The computation time for finding a feasible path is 20 minutes. The minimum distance, in terms of linear movements, from pick to place location is 58 units. CGA subjects only two strings i.e. one fittest string and one offspring for collision computation in the successive generations except the initial fitness ranking generation, thus saving a large amount of time.

The pictorial view showing the object position at intermediate locations for A*, GA and CGA are shown in figure 7, figure 8 and figure 9. These discrete object positions are drawn by converting the intermediate configuration angles of the cooperative manipulator system in C-space to the Cartesian Space.

From the figure 7, it is observed that the discrete object positions are placed in a zigzag position in A*. This is due to the move taken by the search for adjacent feasible configuration when it encounters one. In the Figure 8, it is observed that discrete object positions computed with GA are not located at equal intervals. Since the entire population undergoes cross over and mutation at the same time, the possibility of a larger random joint angle movement exists. In the Figure 9, it is observed that discrete object positions computed with CGA are almost located at equal intervals. Two populations i.e. highly biased individuals from string population undergo adaptive cross over and parameter based mutation, which results in the possibility of smooth movement of joints with equal interval of displacement.

5. CONCLUSIONS

This paper presents a new approach using CGA for automated path planning of 2x3 cooperative manipulator system. The suitability of CGA for offline path planning in comparison with other techniques was demonstrated. From the results presented in this paper, the following conclusions can be made:

1. CGA in conjunction with C-space technique proves to be an effective approach to solve path planning problems of cooperative construction manipulators in complex environment

2. Search using CGA is found to be effective approach in terms of computation time, when compared to other search techniques like A* search and GA search.

3. CGA was found to be efficient in generating the shortest path from pick to place location, when compared to other search techniques like A* and GA.

The future work attempts to investigate the suitability of this approach for more complex cooperative manipulator applications like 2x4 i.e. two manipulators each with 4 DOF.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Department of Science and Technology (DST), New Delhi, India, for the financial support given to this research work vide Grant No. III. 5(23)/2000-ET.

6. REFERENCES

[1] Latombe, J.C. (1991) *Robot Motion Planning*, Kluwer Academic Publishers, Norwell, Massachusetts.

[2] Schwartz, J.T. and Sharir, M. (1988) *A Survey of Motion Planning and Related Geometric Algorithms*, Artificial Intelligence, 37, 157-169.

[3] Udupa, S. (1977) *Collision Detection and Avoidance in Computer Controlled Manipulators*, Ph.D. Dissertation, Department of Electrical Engineering, California Institute of Technology, California.

[4] Lozano-Perez, T. (1983) *Spatial Planning: A Configuration Space Approach*, IEEE Transactions on Computers, C-32(2), 108-120.

[5] Khatib, O. (1986). "*Real-Time Obstacle Avoidance for Manipulators and Mobile Robots*," International journal of Robotics Research, Vol. 5, no. 1, pp. 90-98

[6] Schilling, R. J. (1990). *Fundamentals of Robotics: Analysis and control*, Prentice-Hall, Englewood Cliffs, New Jersey

[7] Sun, S., Morris, A.S. and Zalzala, A.M.S. (1996) *Trajectory Planning of Multiple Coordinating Robots using Genetic Algorithms*, Robotica, 14, 227-234.

[8] Rana, A.S. and Zalzala, A.M.S. (1996) *Near Time Optimal Collision Free Motion Planning of Robotic Manipulators using an Evolutionary Algorithm*, Robotica, 14, 621-632.

[9] Davidor, Y. (1991) *Genetic Algorithms and Robotics: A Heuristic Study for Optimization*, World Scientific Publishing Company.

[10] Paredis, J., Westra, R., (1997) *Co-evolutionary Computation for path planning*, Proceedings 5th European Congress on Intelligent and Soft Computing (EUFIT 97), H.-J Zimmermann(ed.), Verlag Mainz, Aachen.

[11] Sivakumar, PL., Varghese, K. and Ramesh Babu, N. (1999) *Path Planning of Construction Manipulators using Genetic Algorithms*, Proceedings of the 16th International Symposium on Automation and Robotics in Construction, Madrid, Spain, 555-560.

[12] Sivakumar, PL., (2001) *Automated Path Planning of cooperative crane lifts using Genetic and heuristic searches*, M.S. Thesis, Department of civil engineering, Indian institute of technology Madras.

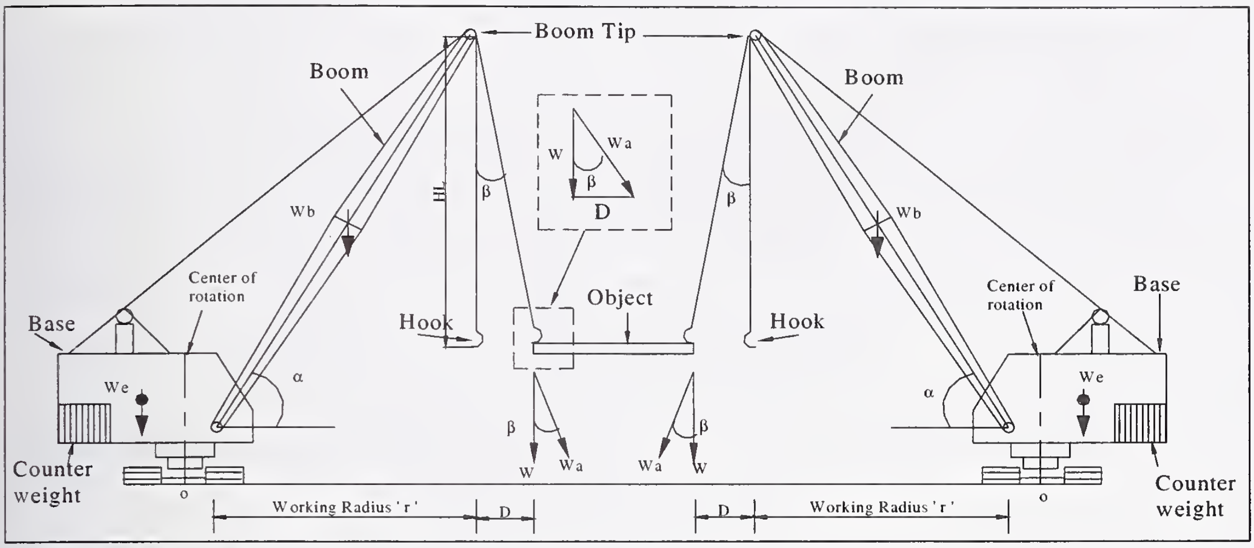


Figure 1 Increase in Load Transferred to the manipulators due to slope of the load line

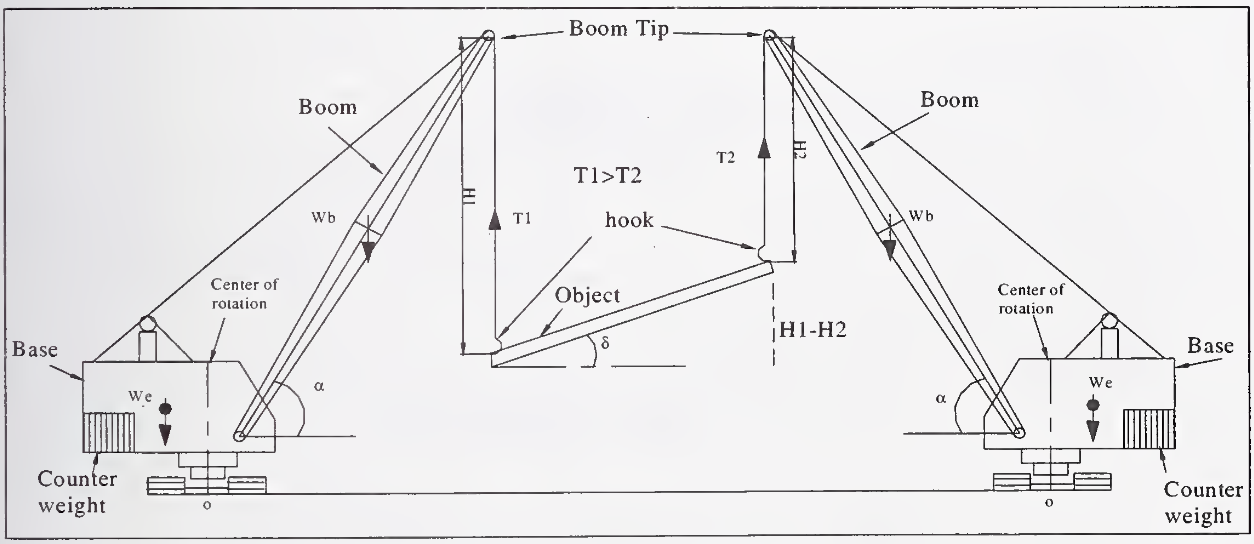


Figure 2 Difference in load shared by hook ends of the manipulators

Table 1. Manipulator arm configurations

Arm configurations	Length	Breadth	Height
Base	10	8	2.5
Boom	40	3	3
Hoisting Rope	39 unit	-	-

Table 2. Upper and lower bounds for the manipulator arms

Arm	Base	Boom	Hoisting Rope
Lower	0^0	30^0	0 unit
Upper	360^0	80^0	39 unit

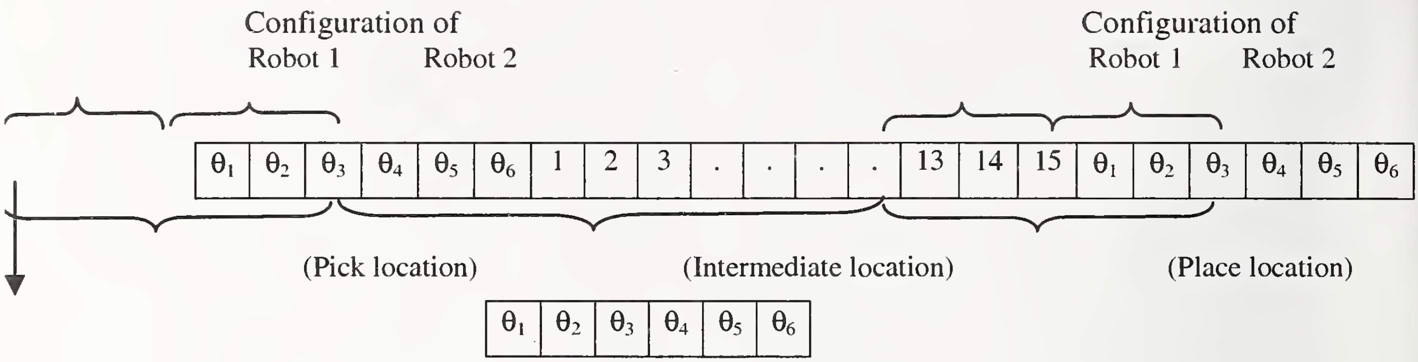


Figure 3.A typical string with fifteen intermediate configurations of 2x3 cooperative manipulators between pick and place location

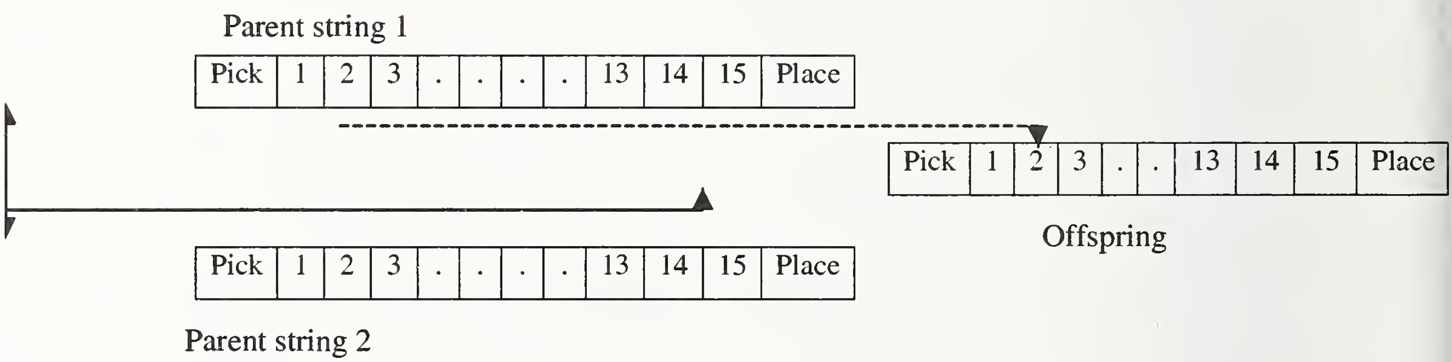


Figure 4. Adaptive cross over

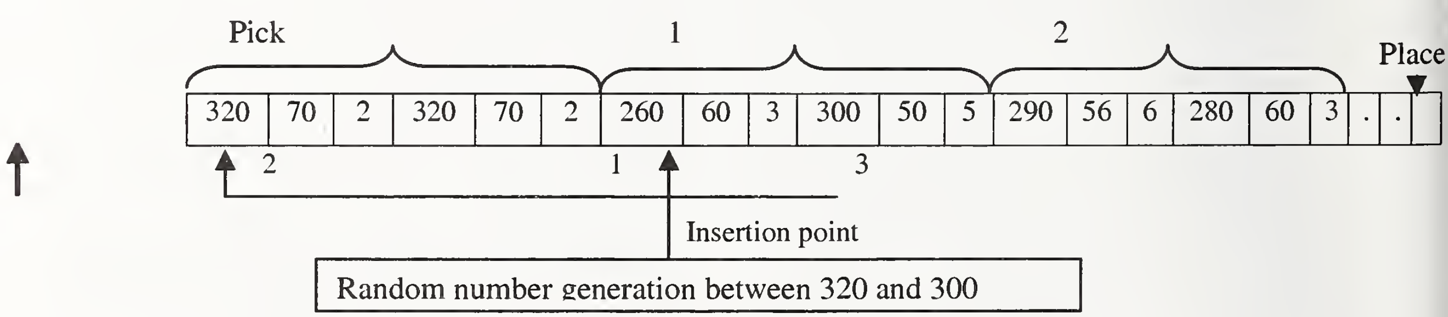


Figure 5. Adaptive mutation

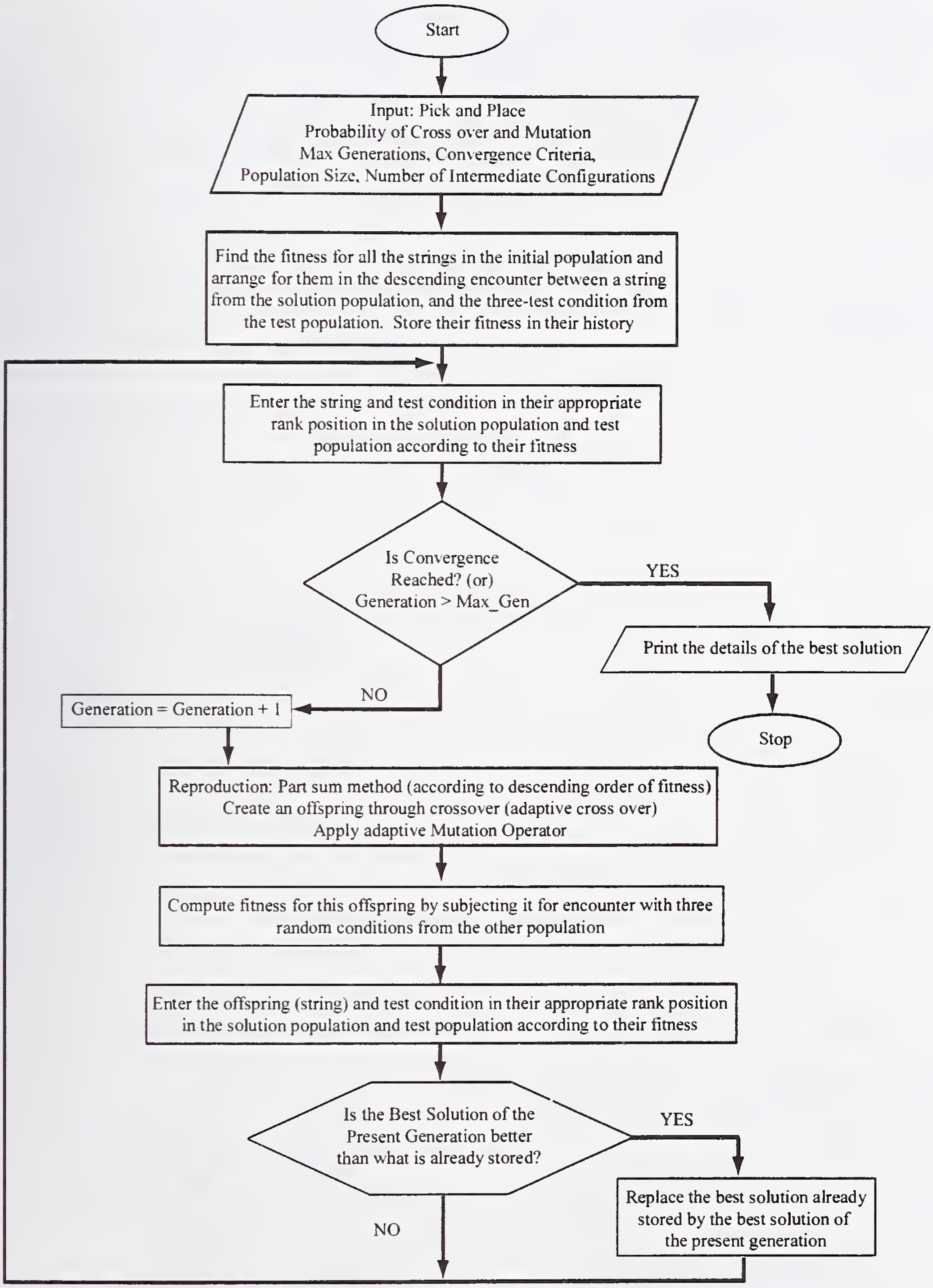


Figure 6: CGA methodology for finding feasible collision-free string

Table 3. Performance comparison of A*, GA and CGA for 2x3 cooperative manipulator path planning

Method	Earlier approach [11,12]		Current approach
	A*	GA	CGA
Pick location	[320,70,2][320,70,2]	[320,70,2][320,70,2]	[320,70,2][320,70,2]
Place location	[270,40,2][270,40,2]	[270,40,2][270,40,2]	[270,40,2][270,40,2]
Number of generation	-	540	750
CPU time (minutes)	320	183	20
Distance in terms of linear movements (units)	73	79	58

Table 4. Incremental angle between adjacent moves of manipulator

Arm	Base	Boom	Hoisting Rope
Lower	0^0	30^0	0 unit
Upper	360^0	80^0	39 unit
Increment	5^0	5^0	1 unit

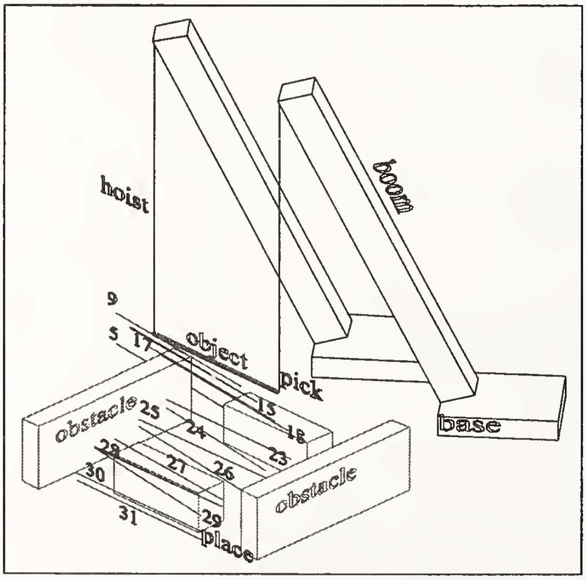


Figure 7. Pictorial view showing the path generated by A* for cooperative manipulators

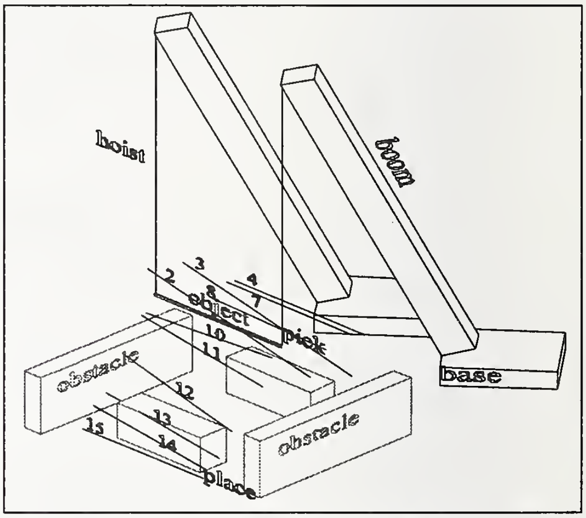


Figure 8. Pictorial view showing the path generated by GA for cooperative manipulators

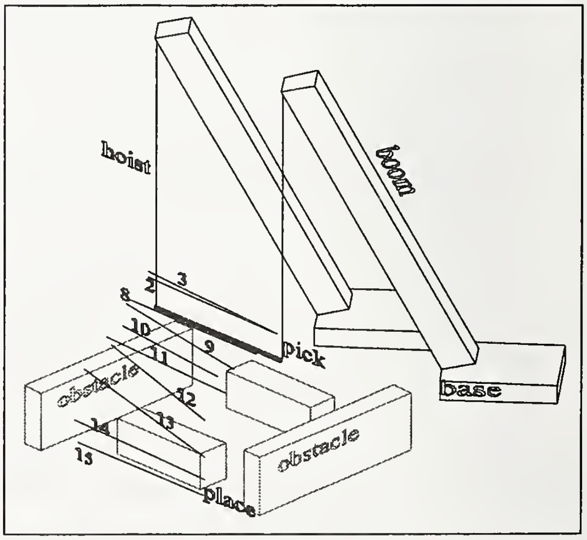


Figure 9. Pictorial view showing the path generated by CGA for cooperative manipulator

MODEL FOR AUTOMATED ROAD-CONSTRUCTION CONTROL

by

R. Navon[1], Y. Shpatnitsky[2] and E. Goldschmidt[3]

ABSTRACT: A survey of current practices revealed that very little control is normally done in earthmoving management. A model to control road construction projects was developed. The model is based on the concept of automatically measuring performance, by measuring locations at regular time intervals, and converting them into controlled parameters. According to this concept, control algorithms convert the measured locations to produce two types of real-time control data: progress and productivity. Locations are measured with a Global Positioning System (GPS). Field experiments conducted with the prototype of the GPS measurement system prove its suitability as a Location Measurement Module.

KEYWORDS: Automated Data Collection; Benchmarking; Control; Earthmoving; GPS; Monitoring; Road Construction.

1. INTRODUCTION

A major obstacle for automated project performance control is measuring the various project performance indicators such as cost, schedule, labor productivity, materials consumption or waste, etc. Advanced technologies, which can be used for on-site measurement and control of project indicators, are emerging and their costs are declining. The work presented here is part of the Technion's initiative called Automated Project Performance Control (APPC). This area broadly refers to the activities taken by the project (or company) management in order to ascertain that the performance of the project is as close as possible to the desirable one. The performance is measured in terms of Project Performance Indicators (PPI). Current efforts focus on automated-data-collection based project performance control, both in building and in earthmoving operations.

The present paper uses a concept of automatically evaluating performance, by measuring indirect parameters and converting them to the controlled variable. The system described herewith will measure the locations, as function of time, of all members of a fleet of earthmoving equipment and convert them to produce real-time control data. The locations will be measured with a Global Positioning System (GPS).

2. INFRASTRUCTURE-CONSTRUCTION CONTROL - STATE-OF-THE-ART

The extent of controlling PPI in infrastructure construction projects is still limited. Current

[1] Assoc. Prof. - Head, Construction Engineering and Management; Head, Construction Automation Laboratory - Faculty of Civil Engineering, National Building Research Institute, Technion City, 32000 Haifa, Israel. ronie@tx.technion.ac.il, http://tx.technion.ac.il/~ronie

[2] M.Sc. Candidate - cvyaniv@tx.technion.ac.il.

[3] Ph.D. Candidate - cveytan@tx.technion.ac.il.

methods normally use manual data collection. The area where Automated Data Collection (ADC) is used for real-time monitoring is quality control and methods that provide data to the equipment's operator, making its manipulation easier.

2.1 Manual Control of Infrastructure Projects

Current practices of project management control in Israel were surveyed in 12 leading construction companies whose main line of business is infrastructure, road construction and earthmoving. The survey was done to assess the state-of-the-art of project management control and the need for it. Most of the surveyed companies admitted that they were not doing any control, but they agreed that if such tools were to exist they would be very willing to use them.

Some of the companies control progress by following up the flow of materials to the site. The bills-of lading are collected and sent to the main office where the quantities are plotted as functions of time and compared to the planned values. The users of this method are generally dissatisfied, reporting that this procedure is inaccurate, error prone (because of the manual data input), and labor intensive. One of the reasons for the inaccuracy is that the planned material consumption assumes theoretical cross-sections, ignoring the difference between the design and the actual physical measurements of the road structure.

2.2 Real-Time Control in Earthmoving Projects

Most previous attempts to automate earthmoving operations are surveying-oriented, yielding impressive results in quantity and quality control. Another aspect of heavy equipment automation is monitoring the operation of the equipment, making its manipulation by the operator easier [e.g. Lee et. al 1997, Peyret et. al 2000, Kannan and Vorster 2000, Bouvet et. al 2001]. The present paper presents the first attempt to automate the control of earthmoving project management.

Lee et al. [1997] proposed a real-time control system to improve the productivity and the quality of asphalt paving operations. The system has four modules using RTK (Real Time Kinematic) GPS. Another model that uses the same technology is called IMPACT [Tserng and Russell 1997]. The model was developed for planning and controlling earthmoving equipment to improve its productivity and safety. The productivity improvement can be achieved by better planning, using simulation. The safety can be enhanced by position measurement, using RTK GPS and controlling the movement to avoid collision.

Recently several other applications of GPS for earth moving equipment control were introduced - most of them with RTK GPS. Peyret et al. [2000] describe a 'Computer Integrated Road Construction' (CIRC) project, aiming at introducing a control and monitoring tool for road pavements construction. The new tools were designed to bring to the sites significant improvements by creating a digital link between the design office and the job site.

Krishnamurthy et al. [1998] describe an automated paving system for asphalt pavement compaction operations. A semi-automated path-planning real-time guidance system that aims towards automating the paving operation was developed. This system accepts relevant paving project inputs, generates appropriate path plans for the compactor, performs a graphical visualization of the generated path plan and offers real-time guidance capabilities using GPS technology.

3. AUTOMATED CONTROL MODEL FOR ROAD CONSTRUCTION

The development of the concept of automatically monitoring performance, by measuring indirect parameters and converting them, was developed in recent years at the Technion [Goldschmidt and Navon 1996 and Navon and Goldschmidt 1999]. This Section sets the theoretical framework of the model and describes the model.

3.1 Conceptual Framework

The system will measure the locations of all members of the fleet of the earthmoving equipment at constant time intervals. An algorithm, described below, will convert these locations to produce two types of real-time control data: progress and productivity. The location will be measured with GPS. The system will use these locations together with data extracted from a Project Model (PM)[4] to determine the activity the equipment is engaged in. its progress, and its productivity. The result will be compared with the planned progress and productivity to give an early warning on deviations as they occur, and will enable the analysis of the causes.

3.2 Model

The model (Fig. 1) compares between the planned and the actual values of progress and productivity variables. The model has two main sources of data: (1) The Project Model, containing the planned schedule, the productivity, and all the data regarding the physical design of the road (layout, cross-sections, etc.). (2) The Location Measurement Module (LMM), using GPS. This module measures the location for each member of the fleet at regular time intervals. The module records the time of measurement, the identification of the equipment and its location.

The model includes four interfaces, which extract all the relevant data from the PM. These interfaces are: Schedule Interface (SI), Geometry Interface (GI), Quantity Interface (QI) and Productivity Interface (PI). The SI begins the process by extracting all the pending activities - these are all the activities whose predecessors are completed, which means that they can be active on the given day.

Specific Work Envelopes (WE)[5] are calculated for each pending activity, based on information in the Knowledge Base, which includes a typical work envelope database, and on the geometrical information extracted by the GI interface from the PM. The WEs correspond with planned work sections, as represented in the schedule. Next, a geometrical calculation associates each of the locations to these specific work envelopes, by checking if the measured location is included within the WE. This, together with Decision Rules from the KB[6], enables the model to determine which activities are actually being performed. Once the model identifies that a new activity has started, it also determines which of the activities are completed. The cycle ends by determining the actual time spent performing each activity, and the productivity, which is based on this time and the completed quantities, extracted by the QI. These data serves as a basis for the output of the model.

The output of the model compares the actual performance, as measured and calculated by the model, to the planned one. It includes:

- A comparison between the actual productivity and the planned one, extracted from the PM by the PI.

- A comparison between the actual progress and the planned one according to the updated schedule extracted by the SI.

The output of the model serves a variety of managerial functions, such as monitoring and taking corrective measures.

---

[4] The Project Model includes a physical description of the road and a description of the activities needed to construct it.

[5] A work envelope is defined to assist the association of the equipment's location to a pending activity, as follows: an area, or volume, where a piece of equipment, working on the road, could be located. The shape and type of a work envelope depends on the nature of the activity, on the construction method, or technology, and the type of equipment.

[6] The decision rules are designed to (1) help associating locations not included in a work envelope, or included in more than one envelope. (2) Determine completed activities.

4. CONCLUDING REMARKS

The first stage of the development was to ascertain that there were no technological barriers. Consequently, a feasibility test of the LMM was carried out. The experiment was planned to check the suitability of GPS, and the pertinent software. The feasibility test is not described here due to space limitations. The experiments confirm that GPS is suitable for the purpose of controlling progress and productivity of earthmoving operations.

The Technion's Automated Project Performance Control (APPC) group is currently engaged in a number of research projects relating to Automated Data Collection. The most notable direction of the group is that of measuring locations at regular time intervals, or other indirect parameters, and using them to automatically control productivity and progress.

5. REFERENCES

Bouvet, D., Froumentin, M., Garcia, G,. (2001). "A Real-Time Localization System for Compactors" Automation in Construction, 10(4): 417-428.

Goldschmidt, E. and Navon, R. (1996). "Automated Real-Time Manpower Productivity Control". The 8th International Symposium on Organization and Management of Construction. Glasgow (UK): 199-205.

Kannan, G., Vorster, M. (2000). "Development of an Experience Database for Truck Loading Opertions." Journal of Construction Engineering and Management, ASCE, 126(3): 201-209.

Krishnamurthy B. K., Tserng H. P., Schmitt R. L., Russell J. S., Bahia H. U. and Hanna A. S. (1998). "AutoPave: Towards an Automated Paving System for Asphalt Pavement Compaction Operations". Automation in Construction, 8(2): 165-180.

Lee, H., Russell, S, H., Schmitt, R, L,. Tserng, H, P., Hanna, A, S., Bahia, H, U., (1997), "Computer Integrated Methodologies for Real Time Control of Asphalt Paving Operations". Proceeding of the 14th International Symposium on Automation and Robotics in Construction (ISARC), 8-11 June, Pittsburgh Pennsylvania: 423-431.

Navon, R. and Goldschmidt, E. (1999). "Automated Labor Control." Proceedings of the Joint Triennial Symposium (CIB Commissions W55 & W65 with the participation of W92) Customer satisfaction: A focus for research & practice, CD-ROM. Cape Town (South Africa), University of Cape Town.

Peyret, F., Jurasz, J., Carrel, A., Zekri, E., Gorham, B. (2000). "The Computer Integrated Road Construction Project" Automation in Construction. 9(5-6): 447-461.

Tserng, H.P. and Russell, J.S., (1997), "Instantaneous Motion Planning and Controlling Tool)IMPACT) for Construction Equipment". Construction Congress Proceedings - ASCE, New York, USA:1020-1029.

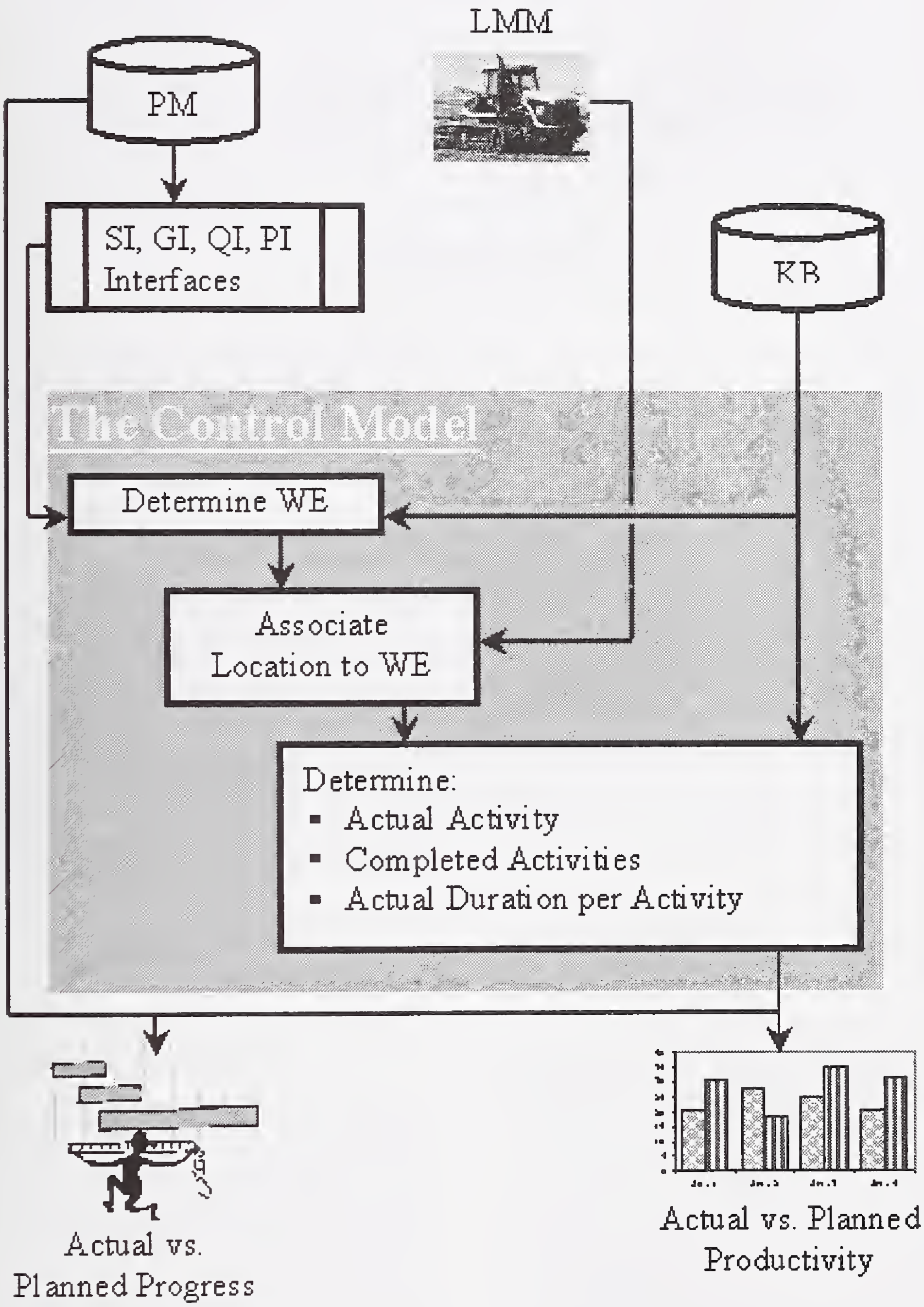


Figure 1: Model Architecture

FORCE SENSOR BY DRIVING HYDRAULIC CYLINDERS: IDENTIFICATION OF INERTIAL PARAMETERS AND TESTS

F. Malaguti, S. Zaghi

Abstract: Force feedback sensors are useful for the planning of robotic digging trajectories. In particular, when combined with force-control algorithms it becomes possible to sense buried objects and to determine the weight of excavated materials. The proposed force sensor system makes use of hydraulic cylinder pressure and thereby measures machine force indirectly. Successful implementation of such an approach will eliminate the need for expensive, direct-force sensors. Measurements of pressures in candidate cylinders were compared with laboratory-measured stroke and force. To measure force in every position and during the motion of a backhoe, kinematic, static and dynamic inertial parameters of the bucket, arm and other links have to be considered. From these data a friction model of the hydraulic cylinders can be developed. The present work involves the determination of these friction parameters for an excavator arm using position measures of the bucket and boom, cylinder pressures and the attitude of the excavator boom. Combined, these represent the "dynamic tare" of system. The work includes a practical approach to filtering regression matrix to do not measure accelerations. The methods allows one to carry out inertial parameters for a full machine, in order to verified design and drawning of costruction machine or to use as input for simulation of machine dynamics.

Keywords: digging robot, force sensor, identification, inertial parameters

Introduction

In previous studies [1,2] static and dynamic analyses of force sensors for digging robots and on-board force-sensing systems for construction machines were presented. The force sensor is based on measurement of pressures in two hydraulic cylinders and angular positions of two links, e.g. with an excavator, the driving cylinders and position of bucket and boom respectively, were measured about an absolute inertial reference point.

Using two cylinders it is possible to compute variable pairs, e.g. force/force-position or force/force-direction, by considering the planar movement of two links. This concept is applicable to many types of construction machines having two or more hydraulic cylinders.

The continuous measurement of soil-bucket interaction force is required to predict interaction force or for real-time planning of bucket trajectories.

Our method of force measurement does not require any additional sensors; it makes use of sensors already available on most construction machinery. It does, however, require new static and inertial parameters for the mechanism links and their respective accelerations and velocities.

Errors in the determination of hydraulic cylinder force by using cylinder pressure measurements are introduced because of the presence of static and viscous frictions. The

CNR-CEMOTER, via Canal Bianco 28, 44044 Ferrara, Italy, email: malaguti@cemoter.bo.cnr.it

proposed new method includes the determination of these parameters.

We use Newton-Euler notation to describe the parameters affecting the dynamic equations of the excavator arm. The parameters are arranged in regression matrix form (eq 1), while the introduction of a linear filter and convolution theorem allowed us to transform the regression matrix into a new system independent of acceleration of the link. This independence avoids the requirement for measuring accelerations of the links (bucket, stick, etc.).

The present work is focused on experimental implementation of this approach, considering practical integration of sensors and software. Tests and experiments were made using a small size excavator (Figure 1) which was modified to use proportional servovalves and was powered by an electric motor. Parameter identification was done off-line using general purpose PC hardare and computing software.

1. Dynamic Analysis and Filtering

Dynamic analysis of the excavator arm shown in Figure 1 was undertaken using Newton-Euler notation.

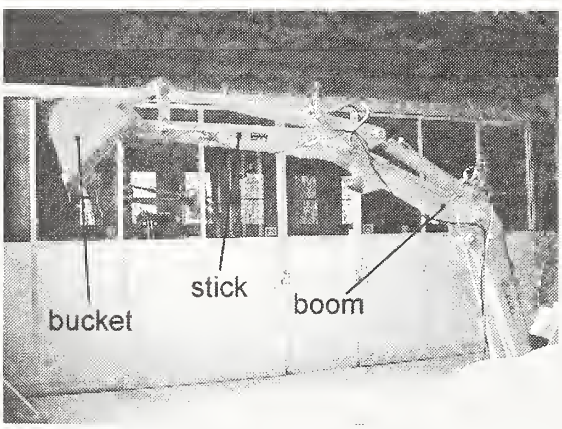


Figure 1 shows the arm of small size excavator modified to fit transducers

The following calculations consider the dynamics of a system comprised of a bucket and stick, without considering the contribution of the boom to the overall excavator dynamics.

Considering the angular positions of the bucket and stick, respectively θ_2 and θ_1 with reference to the horizontal, and the torques trasmitteed by their driving cylinders τ_2 and τ_1, the Newton-Euler dynamics of these links is shown with the system.

$$\begin{bmatrix}\tau_1\\ \tau_2\end{bmatrix}=\begin{bmatrix}I_{11} & I_{12}\\ I_{21} & I_{22}\end{bmatrix}\begin{bmatrix}\ddot{\theta}_1\\ \ddot{\theta}_1+\ddot{\theta}_2\end{bmatrix}+\begin{bmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{bmatrix}\begin{bmatrix}\dot{\theta}_1^2\\ \dot{\theta}_2^2\end{bmatrix} -\begin{bmatrix}h_{11} & h_{12}\\ h_{21} & h_{22}\end{bmatrix}\begin{bmatrix}m_1 g\\ m_2 g\end{bmatrix}+\begin{bmatrix}l_{11} & l_{12}\\ l_{21} & l_{22}\end{bmatrix}\begin{bmatrix}F_x\cos\alpha_x\\ F_x\sin\alpha_x\end{bmatrix} \quad (1)$$

where

$$I_{11}=I_1+m_2r_2l_1\cos(\theta_2+\alpha_2)+m_2l_1^2+m_1r_1^2$$
$$I_{12}=I_2+m_2r_2l_1\cos(\theta_2+\alpha_2)+m_2r_2^2$$
$$I_{21}=m_2r_2l_1\cos(\theta_2+\alpha_2)$$
$$I_{22}=I_2+m_2r_2^2$$

$$h_{11}=r_1\cos(\theta_1+\alpha_1)$$
$$h_{12}=r_2\cos(\theta_1+\theta_2+\alpha_2)+l_1\cos\theta_1$$
$$h_{21}=0$$
$$h_{22}=r_2\cos(\theta_1+\theta_2+\alpha_2)$$
$$C_{11}=m_2r_2l_1\sin(\theta_2+\alpha_2)$$
$$C_{12}=m_2r_2l_1\sin(\theta_2+\alpha_2)$$
$$C_{21}=m_2r_2l_1\sin(\theta_2+\alpha_2)$$
$$C_{22}=0$$
$$l_{11}=l_1\sin\theta_x$$
$$l_{22}=l_1\cos\theta_x+l_x$$
$$l_{21}=0$$
$$l_{22}=l_x$$

and I is the inertia matrix, h is the gravity matrix, F_x is the interaction force on the bucket.

This system was arranged in matrix regression form as:

$$\Delta\tau=W(\ddot{\theta},\dot{\theta},\theta)\Phi \quad (2)$$

where $\Delta\tau=[\tau_2 \quad \tau_1-\tau_2]^T$, and the inertial parameters are defined as (3).

$$\Phi = \begin{bmatrix} I_2 + m_2 r_2^2 \\ I_1 + m_2 l_1^2 + m_1 r_1^2 \\ m_2 r_2 \cos\alpha_2 \\ m_2 r_2 \sin\alpha_2 \\ m_2 l_1 + m_1 r_1 \cos(\alpha_1) \\ m_1 r_1 \sin(\alpha_1) \end{bmatrix} \tag{3}$$

To solve equation system (2) without the measurement of angular accelerations of links, we had to filter equation (1) linearly. By convolution, the filtered torques become

$$\tau_f = \dot{f} * h + f(0)h - fh(0) + f * g$$

were f is the linear stable filter, * is the sign of convolution,

$$h = \frac{d}{dt}\left(I(\theta)\dot{\theta}\right)$$

$$g = -\dot{I}(\theta)\dot{\theta} + C(\dot{\theta},\theta)\dot{\theta} + G(\theta)$$

In this way a filtered regression form of the arm dynamics can be written as:

$$\Delta\tau_f = W_f(\dot{\theta},\theta)\Phi \tag{4}$$

where the vector of parameters Φ does not change, since the filter is linear.

Until now we have used the "torque" term, but it is produced by the cylinder force and crank length. The torques are related to hydraulic pressures of the cylinder by the following equation

$$\tau = FL\sin\left(\arctan\left(\frac{f\sin\theta'}{L - f\cos\theta'}\right)\right) \tag{5}$$

where L is the crank length, f the distance between the link and cylinder pins, and θ' is the angle between those two. This relationship shows that the torque depends on the angle θ' (fig.2), which is related to the θ_i angles of links.

However, it is also possible to use a polynomial approximation of $d\theta_i/dL_i$, which represents the relationship between the angles θ_i and the rod extensions L_i.

The friction of the hydraulic cylinder is difficult to model and compute, because it depends on many factors, for example the types of seal and load pressure. However the same method of identification of inertial parameters allows for the contribution of friction too.

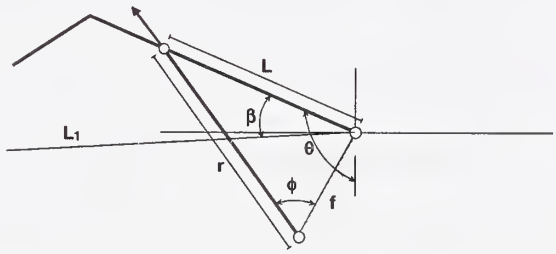


Figure 2: Typical example of mechanism driven by hydraulic cylinder, where r is the variable length of cylinder, L and f the fixed lengths.

From traditional Coulomb friction theory

$$F_f = f_c \,\mathrm{sgn}(v) + f_v v \tag{6}$$

where v is the velocity of the rod, we have two friction coefficients, the f_c Coulomb friction coefficient and the f_v viscous friction coefficient. Remembering that $\tau d\theta = Fdl$, the previous coefficients form the vector of friction parameters :

$$\Phi_f = [f_{c2} \quad f_{v2} \quad f_{c1} \quad f_{v1}]^T \tag{7}$$

2. Inplementation

To filter the regression matrix a simple stable first order filter was used, having the simple transfer function $f(s) = 1/s + 1$ and the following impulse response in the time domain $f(t) = e^{-t}$.

The regression matrix was convolved by this filter, obtaining the filtered regression matrix, independent of accelerations.

In practice we chose the following transfer function:

$$F(s) = a/s + a,$$

with a cutting frequency of 20 Hz, a sample frequency of 100 Hz and $a \cong 125\ rad/s$.

The identification process was based on Least Square Methods considering the measured noise of torque, ε, only.

$$\Delta\tau_f = W_f(\dot{\theta}, \theta)\Phi + \varepsilon$$

Moreover, the components of the torque vector

$$\Delta\tau = [\tau_2 \quad \tau_1 - \tau_2]^T$$

are decoupled, so every link or equation can be considerd as a single input, single output system to estimate.

The measurement hardware consisted of traditional strain gauges and high pressure transducers. These were used to measure hydraulic pressure in both chambers of the cylinders, two resistive angular position trasducers and a tilt sensor.

One position sensor was placed on the stick-boom pin, the other one was placed on the pin of the crankshaft that transformed the linear motion of rod into angular motion (see fig.1). The angular position sensor location was chosen for practical reasons, so that the relationship between the force produced by the cylinder and the angular position of the bucket was known, based on the kinematics of the bucket mechanism.

The tilt sensor (Accustar) placed on the boom of our small excavator provided the absolute inertial reference to measure the angular positions of links. Additionally it allowed the determination of the boom angular position and velocity.

The signals were acquired by a standard multifunction National Instrument I/O board for PCs and processed using Matlab-Simulink. The same software was used for off-line system identification routines.

As discussed previously, the solution to filter dynamic system and its regression form was chosen to avoid the use of acceleration sensors or the noisy and inaccurate double time derivative of angular displacement. However the regression form needs to measure or estimate angular velocity.

Because the simple method of backward differencing does not yield good performance, the angular velocities were estimated using the Savitzky-Golay filter.

In practice this filter, performing a least squares linear regression fit, is a mobile window of *2nw+1* number of samples of angular position that is fitted with an m-order polynomial.

$$\theta(t) = p_1 t^m + p_2 t^{m-1} + \cdots + p_m$$

The time derivative or angular velocity is derived at the same time. To avoid high-order calculus the filter parameters were chosen as: *nW=3* and *m=2*. The velocity is thus given as:

$$v = 2p_1 t + p(2)\ .$$

Also in this case the filtering was done off-line.

First, static inertial parameters were determined in the static configuration; afterward the fitting of all static, dynamic and friction parameters was conducted during dynamic conditions, that is, by moving the arm of the excavator.

In the first case we used the normal regression matrix because the static conditions do not depend on accelerations, while under dynamic conditions the filtered regression matrix was used.

Many references on identification of dynamic parameters of industrial robots suggest the use of specified trajectories to minimize the conditioning of the data, but in our case it was impossible to plan suitable trajectories that properly simulate a digging cycle under manual control.

3. Results of estimation

The tests to estimate parameters of the excavator arm can be divided in two parts: inertial static parameters in static condions, and all inertial and friction parameters in dynamic conditions.

3.1 Static estimation

Static estimation was carried out by putting the excavator arm, and in particular the stick and bucket, in various positions, where angular position and torques for each joint was measured with acceleration and velocity of the links equal zero.

This set of measurements was repeated for a number about 150 samples (positions), to have small variation in the results. We avoided those positions corresponding or close to the endstop of the rods.

In this first part we chose to not include the static friction parameters in the estimation, because the static friction depends on many factors including the characteristics of the hydraulic fluid, the type of seal, and the time that the rod remains in the same position, making difficult to have a reliable model of friction.

The results of the static estimation are summarized in Table 1,

i	ϕ_i [kgm]	$\hat{\sigma}_i$ [kgm]
S1	6.17	2.95
S2	12.05	4.45
S3	103.31	23.67
S4	22.78	7.08

Table 1: Parameter estimation of static conditions

where Φ_i and σ_i are the vectors of static parameters and their standard deviations respectively.

3.2 Dynamic estimation

For the estimation of all parameters in dynamic conditions, the signals were sampled at 100 Hz for 40 seconds, converted and processed, to bc estimated off-line as described previously.

The final ersults are shown in Table 2:

parameters	Estimated value	Standard deviation
ϕ_{D1}	100.32 kgm^2	25.7 kgm^2
ϕ_{D2}	186.73 kgm^2	33.4 kgm^2
ϕ_{S1}	3.50 kgm	1.54 kgm
ϕ_{S2}	13.71 kgm	3.28 kgm
ϕ_{S4}	71.81 kgm	10.07 kgm
ϕ_{S4}	35.56 kgm	5.68 kgm
ϕ_{f1}	52.23 kgm/s^2	25.5 kgm/s^2
ϕ_{f2}	31444.87 kg/s	1257 km/s
ϕ_{f3}	35.67 kgm/s^2	20.6 kgm/s^2
ϕ_{f4}	30167 kg/s	1320 kg/s

Table 2: Parameter estimation in dynamic conditions

One way to value the quality of estimation is to compare the measured and estimated torque data. Figure 3 shows the measured and estimated torques related to the bucket, while Figure 4 shows a comparison of stick torques. By analyzing the residual vector, the difference between measured and predicted torques were computed. The mean square error and percentage error are given in Table 3.

torque	Mean square error	Percentage error
τ_2	43.12 Nm	26.7 %
τ_1	73.54 Nm	18 %

Table 3: Mean Square and percentage eerors of stick torques.

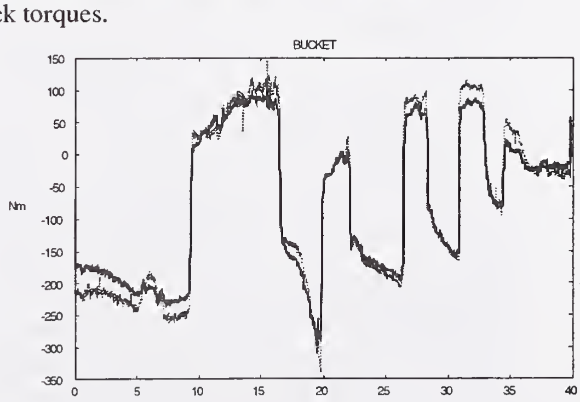


Figura 3 Comparison between measured (full line) and estimated (dashed) torques of bucket

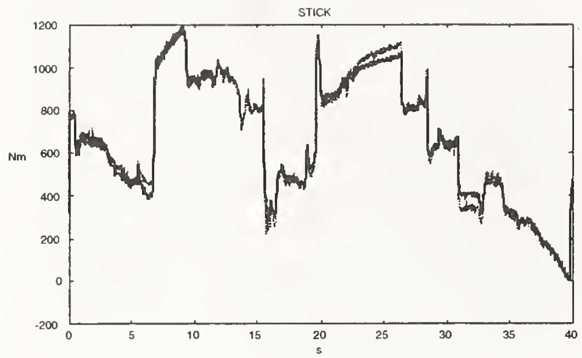


Figura 4 Comparison between measured (full line) and estimated (dashed) torques of stick

As we can note, the torque of the bucket shows a larger error with respect to the stick torque. This is probably caused by many factors: the bucket motion needs only small forces and low pressures, so the influence of the joint and rod seal frictions will have more uncertain values and the variation is therefore higher; second, the angular position is measured indirectly. Also, the full scale range of the pressure sensors utilized was high in comparison to the pressures needed to actuate the bucket.

4. Conclusions

This work purses two fundamental objectives: first to develop an economical force sensor for contruction machinery; and second, to develop a method to measure the masses, moments of inertia, and friction of parts, when it is not easy to compute these parameters, or to validate models by running mechanical dynamic simulation software.

To achieve these objectives, a small excavator, powered by electric motor and controlled by servo valves, was equiped with a suitable sensor suite. For machine dynamics, a filtered regression approach was studied.

In these first tests we can see that the parameters indentified in dynamic conditions and considering cylinder friction are more accurate with respect to static conditions.

We can note that the estimated torques follow the trend of measured torques, but it is evident that the errors on the torque related to the bucket are large with respect to the error of stick torque.

We were unable during the course of the research to find the reasons for these discrepancies. These parameter values depend on a number of things, from selecting the appropriate pressure sensor range to software routines, hydraulic system configuration, and signal processing techniques.

This complexity requires an accurate calibration of all these factors, and that takes time. The initial results suggest that with further work this technique can be made to work, thus providing an economical sensor approach for control of machinery.

REFERENCES

[1] F. Malaguti, "Development of On-board Weighting and Force Sensor System for Construction Machines: Static Analysis", Proceedings of 14th ISARC, Pittsburgh, USA, June 8-11, 1997.

[2] F. Malaguti, "Force Sensor for Construction Machines by They Driving Hydraulic Cylinders", Proceedings of 17th ISARC, Taipei, Taiwan, 18-20 September, 2000.

[3] A.I. Koivo, "Fundamentals for Control of Robotic Manipulators", J. Wiley & Sons., 1989

[4] F.L. Lewis, C.T. Abdallah, D.M. Dawson, "Control of Robot Manipulators" Macmillan, New York, 1993.

[5] S. Tafazoli, P.D. Lawrence, S.E. Salcudean, D. Chan, S. Bachmann, C:W. De silva, "Parameter Estimation and Actuator Friction Analysis for Mini Excavator"

[6] S. Tafazoli, P.D. Lawrence, S.E. Salcudean "Identification of Inertial and Friction Parameters for Excavator Arm", IEEE Trans. on Robotics and Automation, vol. 15, no. 5, October 1999.

INTELLIGENT DRIVING CONTROL SYSTEM

-Automated Driving Management and Control System Using Optimum Control Theory-

Teruhisa AWATA (awata@ansl.ntt.co.jp)
Koichi YOSHIDA (yoshida@ansl.ntt.co.jp)
Hidenori HINO (hino@ansl.ntt.co.jp)

NIPPON TELEGRAPH AND TELEPHONE CORPORATION
Access Network Service Systems Laboratories

1-7-1 Hanabatake, Tsukuba-city, Ibaraki, 305-0805, JAPAN

Abstract: To realize the all NO-DIG (trenchless method) operation of conduit installation, NTT developed the microtunneling method "ACEMOLE" allowing long-distance and curved installations and promoted use of this method. NO-DIG method is not widespread in construction works because its construction cost is not economically compared with open-cut method. We developed new ACEMOLE to solve these problems drastically and realize NO-DIG entirely. Especially position-detection system using fiber-optic gyroscope and intelligent driving control system using optimum control theory make possible to control microtunneling machines accurately without operators skills and to disseminate NO-DIG technology smoothly.

Keywords: NO-DIG(trenchless method), microtunneling robot, fiber optic gyroscope, directional control, Kalman filter

1 Introduction

NTT has independently developed and promoted the microtunneling robot 'ACEMOLE' as a technology for laying electronic communications conduits. The application of the ACEMOLE has expanded to take in other pipe and conduit-laying operations, such as water and sewage pipes and electric power cables. This technology has now laid 677 km of conduit, but costs are high, numerous technological problems have been encountered and the level of NO-DIG operations have reached their limit. Cost reduction is the major prerequisite for increasing the application of this NO-DIG method. Another important agenda is the improvement of microtunneling machine position detection and control technology, to enable flexible response to varying road shapes, operation in areas with a high concentration of underground facilities and high-accuracy long-distance and curved route driving. Accurate control of the microtunneling machine requires a system which accurately monitors the position and posture of the machine, predicts changes in driving status and machine behavior with changes in external conditions, and determines control guidelines on the basis of an overall judgment of status. However, depending on the form of the route being driven and the driving distance, there are certain restrictions on present position detection technologies. Problems may occur with the accuracy and continuity of measurements, or measurements may not be able to be taken in real time. In addition, human error on the part of the operator or the application of control with an inaccurate understanding of changes in external conditions can cause trouble, and this trouble can spread without being corrected. It is necessary to prevent or reduce the occurrence of trouble originating in human error and external factors to achieve the smooth spread of NO-DIG method.

NTT has therefore developed continuous high-precision position detection system to enable real time position detection without restrictions over the entire driving course and driving distance using optic fiber gyroscopes. In addition, to further prevent the occurrence of trouble and enable higher-accuracy execution, an Auto-navigation system has also been developed. The introduction of these technologies enables inexperienced operators to control the microtunneling machine at the same level as experienced operators, prevents the occurrence of various types of trouble which have occurred in execution up to the present, and has improved the accuracy and efficiency of driving operations. This paper will give an outline of the overall makeup of the ACEMOLE system, and will then discuss the high-precision continuous position detection system. After this, the makeup and component technologies of the Auto-navigation system will be discussed. The paper will conclude with a discussion of the results of test operations.

2 Auto-navigation System

2.1 Outline of ACEMOLE System

Fig. 1 shows the entire system configuration of the microtunneling ACEMOLE robot controlled by the system described in this paper. The ACEMOLE system shown here uses dynamic press-insertion, in which vibration of the front of the microtunneling machine liquefies the surrounding soil face to reduce face resistance and the machine advances under thrust provided by jacking machine. This system lays pipes of a fixed length in succession. The machine's driving direction is controlled by an inbuilt hydraulic jack, which moves the front of the machine head up and down and left and right.

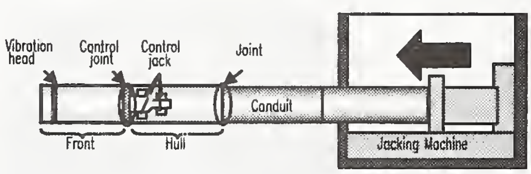


Figure 1: Schematic of a microtunneling robot

2.2 Position Detection System

Fig. 2 shows the horizontal and vertical position detection systems. Up to the present, horizontal position detection has relied on a laser target system: a laser beam produced by a laser theodolite is detected by a light-sensitive target in the microtunneling machine, and the displacement of that beam from the base line (laser beam) is determined. An Electromagnetic method is also used, in which receivers above ground detect a magnetic field induced by a transmission coil built into the microtunneling machine, enabling the machine's absolute position to be determined. However, the laser targeting method can only be used over straight sections, and the magnetic method used on curved routes requires the operator to stop driving in order to take above ground measurements, resulting in discontinuous position data. The magnetic method also cannot be used near magnetic bodies, under riverbeds, etc. An optic fiber gyroscope system was developed to provide a solution to these problems. Vertical position is detected using a Hydraulic pressure differential method enabling relative depth to be detected in real time, using the difference between a standard liquid pressure and the pressure detected by a pressure sensor in the microtunneling machine.

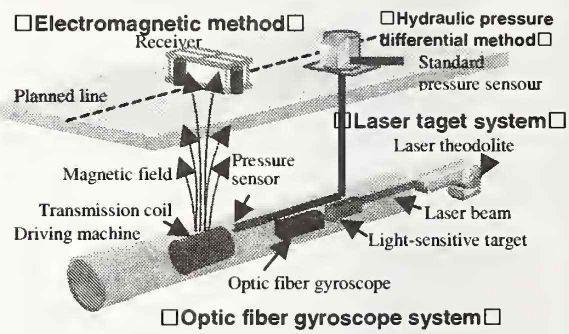


Figure 2:.Position detecting system

2.3 Outline of Optic Fiber Gyroscape System

Optic fiber gyroscopes have a broad range of application, but are chiefly utilized in the inertial measurement of direction and posture in space and the measurement and control of rotation. These gyroscopes were utilized in a newly developed system for microtunneling machine position detection as a means of solving the problems delineated above. Optic fiber gyroscopes continuously detect the angular velocity of a moving body. relative change in the angle can be found by integrating the angular velocity, and therefore the horizontal position of the machine can be continuously calculated from the driving distance using the starting point of driving as a base. Horizontal position y is found using the following equation, where Ω is angular velocity and x is driving distance:

$$y(x) = \int_x \int \Omega dt dx \quad (1)$$

Below is a simple discussion of the Sagnac effect (the principle of measurement with optic fiber gyroscopes). We will consider the optic system shown in Fig. 3. The beam of light produced by the light source is split into left and right beams by a beam splitter. These beams then enter circular optic fiber loops. After being propagated through these loops, the beams are reintegrated at the same beam splitter. If the loops have remained still, the left and right beams will have traveled a path of exactly the same distance, and the time required for the beams to arrive at the beam splitter will be the same. However, if the optic fiber loops are rotated in one direction at angular velocity Ω, then the distance traveled by the two beams will differ, and there will be a difference in propagation times. This time differential is expressed as

$$\Delta t = \frac{2rL}{c^2}\Omega$$

where L is the length of the fiber, r is the radius of the fiber, and c is the speed of light.

(2)

The time differential Δt for propagation of the left and right beams is proportional to the angular velocity of the rotation of the optic fiber loops, Ω. Calculation of angular velocity utilizes the fact that the propagation time differential for the left and right optic fiber loops Δt found in equation (2) becomes $\Delta\theta$, the phase difference between the left and right beams, as shown in equation (3).

$$\Delta\theta = \omega\Delta t = (2\pi / \lambda)c\Delta t$$

ü@ü@ü@ü@ü@ $\frac{4\pi rL}{c\lambda}\Omega$ (3)

Here, ω is the angular frequency of the light source.

Utilizing equation (3) to find phase difference $\Delta\theta$ enables the rotational angular velocity Ω of the fiber optic loop to be determined. In addition to the use of angular velocity Ω when calculating horizontal position, this parameter is also used as important machine posture data in the Auto-navigation system described in the next section.

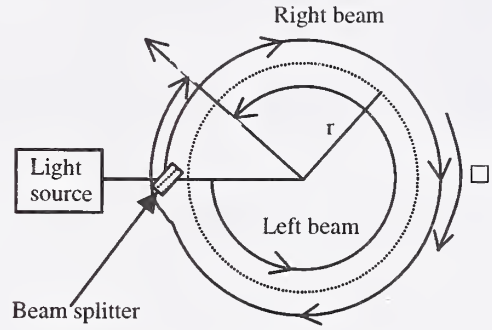


Figure 3:.Principle of Sagnac effect

2.4 Outline of Auto-navigation System

The Auto-navigation system detects the machine's position and posture, predicts behavior and provides control. Fig. 4 shows an image from the Auto-navigation system screen. The screen shows 0 m for both horizontal and vertical position. On the left (the negative side), we see data detected to present and values predicted on the basis of the machine behavior model (to be discussed in the next section). On the right (the positive side), we see the optimum value of correction of machine direction to be applied based on model parameters estimated to present, and results of a simulation of machine displacement in response to this correction.
Accurate detection of microtunneling machine position and posture are essential for directional control of the ACEMOLE. This is an important element in the application of control and has an effect on driving accuracy. The optic fiber gyroscope system described in Section 2.3 is able to gather high-precision position data in real time; conditions in the machine behavior model can therefore be accurately estimated online, and the optimal value of control can be determined by feedback of these conditions, enabling high-accuracy directional control.

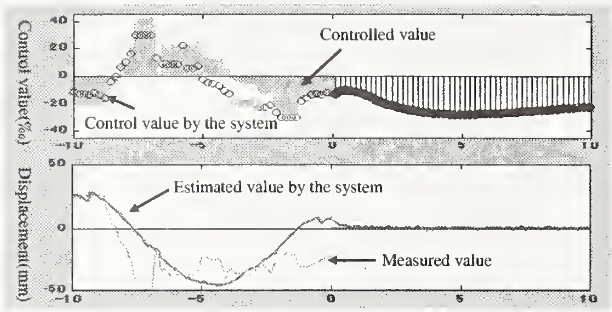


Figure 4:.Screen shot

2.5 Machine Behavior Model

Because horizontal and vertical motion is independent and mutual interference does not occur, models of each direction have been designed. Because the characteristics of the machine's motion in the horizontal and vertical directions can be expressed in identical behavior models, this model will be explained using Fig. 5, which defines machine position and posture in relation to the microtunneling machine base line.

Changes in machine posture ($\Delta\theta$) are defined by the linear sum of noise components which are unrelated to past head angle (including current head angle) and directional control. This can be expressed as follows:

$$\theta[k+1]=\theta[k]+\sum_{i=0}^{nb-1} b_i \eta[k+1-i]+\omega[k] \quad (4)$$

Here, k is the positive increase in unit driving length L_p for each length driven, $\eta[k+1-i]$ is L_p*i(m), value of past directional control conducted before the present, b_i is the weighting factor for value of directional control, and ω represents noise components.

If the angle of the rear of the microtunneling machine as measured against the base line is taken as following the angle of the front of the microtunneling machine at L_f(m) only, it can be expressed as the following:

$$\theta_r[k]=\theta[k-nd] \quad (5)$$

where nd is L_f/L_p rounded to a whole number.

The change in displacement X_h of the section of the microtunneling machine receiving directional control for each driving length can be expressed

$$X_h[k+1]=X_h[k]+L_p\theta[k] \quad (6)$$

Until now, it has been impossible to directly detect θ in the horizontal direction and it has been necessary to use a kalman filter adapted to the machine posture model given as equation (4) and the position and posture detection method being used to estimate X_h and b_i from equation (4). However, because the optic fiber gyroscope system has enabled direct detection, it has now become possible to construct a more accurate model.

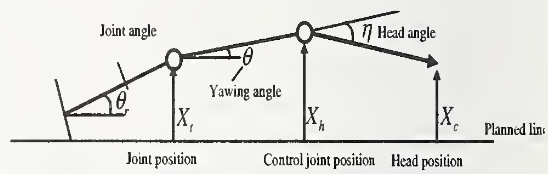


Figure 5:.Coordinate system for the driving machine

2.6 Optimum Feedback Control

We will now consider directional control conducted using the equations described above. Equations (4), (5) and (6) are used to newly define the posture vector X.

$$X[k]=\begin{bmatrix} \theta[k] \\ \theta[k-1] \\ \vdots \\ \theta[k-nd] \\ \cdots\cdots \\ \eta[k-1] \\ \vdots \\ \eta[k-nb+1] \end{bmatrix} \quad (7)$$

nb in equation (7) is an integer greater than 2. Next, a feedback gain K which satisfies equation (8) is sought.

$$\eta[k+1]=\eta[k]-\boldsymbol{KX}[k] \quad (8)$$

K is designed on the basis of the principle of harmonizing speed of convergence with value of correction. This has made it possible to closely control the progress of the microtunneling machine underground, and to rapidly correct deviations from the base line.

2.7 Auto-navigation System Flow

Fig. 6 shows the composition of the Auto-navigation system. The utilization of a model of microtunneling machine behavior (Fm) which expresses the value of directional control η (k) and the motion characteristics of the microtunneling machine in response to changes in position and posture mean that current machine status can be estimated online. Feedback of these estimates enables the optimal directional correction controller (K) to calculate the value of correction η (k+1) to be applied next. This online estimation enables the appropriate value of control to be exerted to maintain the microtunneling machine head on the planned line in response to current position and posture, surrounding soil quality, etc., and future machine status to be predicted.

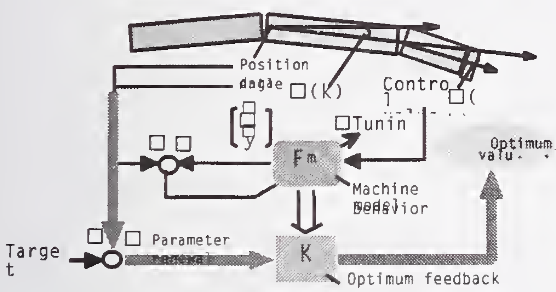


Figure 6:.Composition of Auto-navigation system

3 Test Results

To verify the effectiveness of the system, proving trials were conducted using an actual machine. The machine was placed in a homogeneous stratum of 3 m in width and 40 m in length, and driving commenced from a depth of 2 m. Figs. 7 and 8 show actually measured values and values estimated by the system for horizontal and vertical directions respectively. The horizontal axes show data sampled at L_p = 10 cm.

Values predicted by the system are compared with values measured in the horizontal direction using the magnetic method (capable of measuring absolute position in this direction) and in the vertical direction by the liquid pressure differential method. The results of directional control conducted on the basis of optimum values calculated by the system showed less than ± 30mm difference between values predicted by the system and actually measured values, proving the effectiveness of the model and the algorithm.

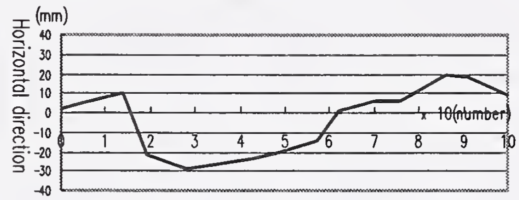


Figure 7:.Values estimated by the system and measured values by electromagnetic method

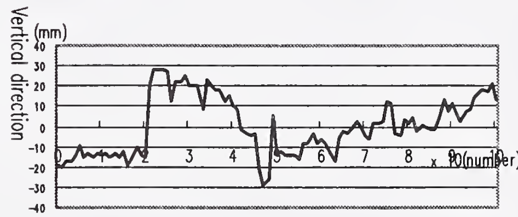


Figure 8:.Values estimated by the system and measured values by Hydraulic differential

4 Conclusions

A high-accuracy Auto-navigation system has been developed for conduit microtunneling machines using the dynamic press-insertion method. The system utilizes a machine behavior model which displays motion characteristics in response to correction of microtunneling machine direction. This has enabled optimal feedback control based on the model and estimates of position and posture of the microtunneling machine, both of which are difficult to directly monitor. Use of an optic fiber gyroscope for horizontal position and posture detection has enabled estimates of machine status to be made at a higher level of accuracy. The introduction of this system has increased the efficiency of execution management and optimized driving control, leading to the following benefits:

(1) Homogeneous and high-accuracy execution is possible irrespective of the level of experience and skill of the operator.

(2) Rapid and accurate driving control has increased the level of driving efficiency achieved to date.

(3) The problem of shortage of skilled operators has been solved, and the period

and costs required to train operators have been reduced.

(4) These results will aid in the diffusion of NO-DIG method.

The results reported in this paper have demonstrated the effectiveness of the newly developed technology. To increase the spread of NO-DIG method in the future, it will be essential to continue to improve the technology in this way. These technologies have been developed for the ACEMOLE system, but the broad applicability and potential for future development of the system mean that we can expect that it will be possible to construct similar systems reflecting machine characteristics for other driving methods.

5 REFFERENCES

(1) Ayukawa, M. Inoue, K., Miyatake, M., and Tujimura, T., "The Future ACEMOLE Realizes All No-Dig Operations with Conduit Installation," NTT Review Vol. 13, No. 5, pp 69-77, 2001.

(2) Manabe, T., Sato, M., Awata, T., and Matsumoto, M., "Automatic Direction Control Technique for Press-in Type Small Diameter Tunneling Machine, " ISTT International Conference on No-Dig '98," pp. 83-92, 1998.

(3) Awata, T., "Intelligent Driving Control System," No-Dig International, Vol. 12, No. 4, pp 25-26, 2001.

(4) Maybeck, P., "Stochastic Models, Estimation and Control," New York Academic, 1982, Vol. 1.

AUTOMATIC STEERING SYSTEM FOR ROTARY SNOW REMOVERS

Hirofumi HIRASHITA, Takeshi ARAI, Tadashi YOSHIDA

Advanced Technology Research Team, Public Works Research Institute
1-6,Minamihara, Tsukuba City, Ibaraki-Pref, Japan, 305-8516
Tel:+81-298-79-6757 Fax:+81-298-79-6732
E-mail:hirasita@pwri.go.jp, arai0333@pwri.go.jp, yoshida@pwri.go.jp

ABSTRACT: In cold and snowy parts of Japan, most snow is removed from roads by mechanical snow removers, and to deal with roadside conditions and guarantee safe operation, these machines are normally operated by two people: an operator and an assistant.
Work is in progress to develop automatic steering systems for snow removers in order to lighten the burden on snow remover operating staff and to reduce future labor requirements. This report describes an automatic steering system that incorporates three technologies: the lane marker system that has been developed in recent years as an ITS technology, RTK-GPS technology, and GIS technology. This steering system has been developed for rotary snow removers: the type of snow remover considered to be the most difficult to operate. The report also evaluates the three control methods based on the results of corroborative testing done using actual snow removers and outlines problems with these control systems that must be overcome to establish a working system.

KEYWORDS: Rotary snow remover, automatic steering, ITS, lane marker, GIS, GPS

1. INTRODUCTION

In cold snowy parts of Japan, snow is removed from roads by mechanical means to guarantee smooth winter road traffic. To guarantee that mechanical snow removal is performed appropriately according to the state of snow accumulation on road surfaces, a carefully planned attendance control method must be established and the machinery must be operated correctly.

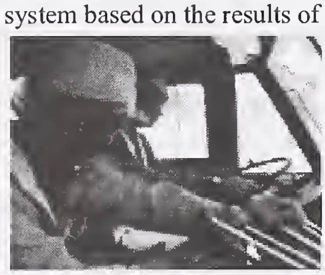
The automatic steering system that has been developed is intended for use on rotary snow removers: a type with an operating method more complex than that of other types of snow removers. It has been developed to reduce the burden on snow remover operators and to lower future labor requirements.

2. BACKGROUND TO THE DEVELOPMENT

A rotary snow remover is operated by two people: an operator who drives the remover and an assistant who controls the ejection of the snow. The operator drives the machine along the curb on the shoulder of the road while watching out for manholes or other level differences on its surface. The assistant ejects the snow while avoiding private property, homes, and other areas where snow disposal is forbidden (Photo 1).

The goal is to automate part of the vehicle operating work now performed by the operators to allow them to control the snow ejection now done by the assistants, permitting the introduction of one-man operation in the future.

This report introduces positioning technology and control technology that form the foundation of the automatic steering system that has been developed, and presents an evaluation of the system based on the results of corroborative testing performed with an actual snow remover and describes problems to be overcome to complete a working system.



Photograph 1. Inside of a Rotary Snow Remover

3. AUTOMATIC STEERING SYSTEM

The automatic steering system that has been developed automates the rotary snow remover steering operation that is one of the tasks of the operator.

3.1 Steering control of a rotary snow remover

The steering mechanism of a rotary snow remover differs from normal vehicles in that the front and back parts are linked by pins and articulating mechanisms are installed so that the vehicle can bend at the pins. The steering is controlled by the exact linearization method and by time scale transformation that are effective ways to control the course of a moving vehicle [1]. This control method is represented by an equation of motion subject to non-holonomic constraint and is based on non-linear control theory. But a snow remover slides laterally, because as it removes snow, it is subjected to lateral reaction force from the snow embankment as shown by Photograph 2. Sideways sliding is controlled by using the integrated servo system described below to treat this sliding as an external disturbance.

Photograph 2. Rotary Snow Remover Removing Snow

(1) Course tracking control

Fig. 1 shows a model of the snow remover steering mechanism. In this figure, the distances from the connecting pin of the snow remover to the front wheels and to the rear wheels are represented by L. When the steering angle is represented by 2α, the course followed by the snow remover is an arc with radius R = L/tanα. When the course tangent direction speed vector at the center point P of the front wheels is represented by v and the tangent direction angle by θ, the equation of motion of the point P can be written as follows.

$$\dot{x} = v\cos\theta$$
$$\dot{y} = v\sin\theta$$
$$\dot{\theta} = \frac{v}{R} = v\frac{\tan\alpha}{L} \qquad (1)$$

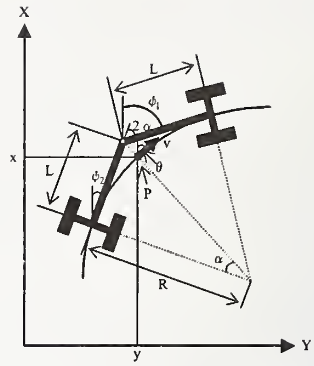


Figure 1. Coordinate System (Model of a Snow Remover)

If exact linearization and time scale transformation are done for equation (1) to perform status feedback control, the quantity of steering of the snow remover that is 1/2 of the steering angle α is defined by the following equation.

$$\alpha = \tan^{-1}\{(f_1 y + f_2 \tan\theta) L\cos^3\theta\} \qquad (2)$$

In equation (2), lateral sliding is not considered. In order to eliminate the steady-state deviation caused by lateral sliding, if the integrated servo system is used, the quantity of steering α is represented by the following equation.

$$\alpha = \tan^{-1}\left[L\cos^3\theta\left\{f_3\int(y - y_{ref})v\cos\theta dt + f_1 y + f_2 \tan\theta\right\}\right] \qquad (3)$$

The symbol f is the control gain, and based on equation (3), f_1 represents proportional gain, f_2 represents derivative gain, and f_3 represents integral gain. PID control was used to construct the system.

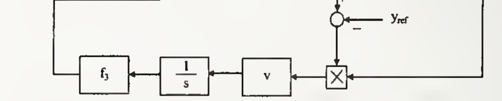


Figure 2. Integrated Servo System Control Block Diagram

3.2 Positioning method

It is necessary to detect a vehicle's present position and to have information that determines the direction to steer it in order to perform automatic steering.

The system that has been developed is equipped with two functions: the guidance method that use lane marker sensors to detect the vehicle's position and the steering method that uses GPS and GIS to detect the vehicle's position.

3.3 Lane marker sensor guidance method

A lane marker sensor is a basic technology of the Advanced Cruise-Assist Highway System (AHS) developed for use as part of the Intelligent Transportation System (ITS). This technology includes lane markers buried at intervals under a road surface and vehicle mounted sensors that detect signals transmitted by these lane markers to guide the vehicle along the course formed by the markers. There are two versions of this system: the radio wave method and the magnetic method.

(1) Steering angle calculation logic

It is assumed that the distance between the lane markers will not be constant, but can be varied according to the shape of the road and demands of the sensor side. Therefore, this system is equipped with two rows of sensors so that control can be performed even when the lane marker installation interval is not known until the rotary snow remover passes over them (Fig. 3).

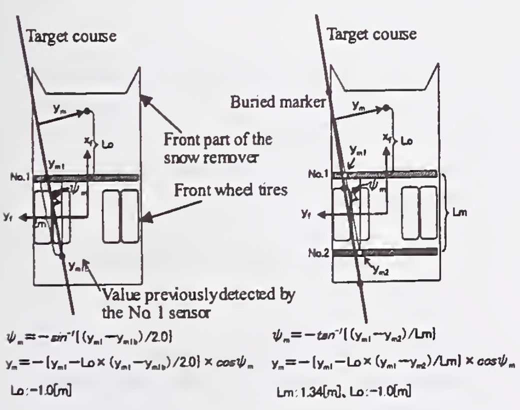
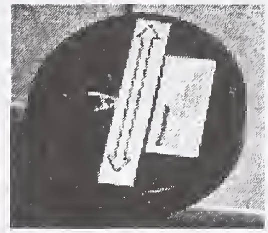


Figure 3. Lane Marker Sensor Guidance Method

To control steering with a single row of sensors (Fig. 3 on the left), the lateral differential and the azimuthal differential y_m and ψ_m that are necessary for control by the target course coordinate system $\sum(xy)$ require information about the interval between the markers in addition to the sensor signal y_{m1} that is measured by the front vehicle fixed coordinate system $\sum f(x_f y_f)$. But with two rows of sensors (Fig. 3 on the right), the system can calculate the lateral differential and the azimuthal differential y_m and ψ_m of the target course coordinate system $\sum(xy)$ by having two opposed sensors hold the two sensor signals y_{m1} and y_{m2} measured by the front vehicle fixed coordinate system $\sum f(x_f y_f)$ as it passed a marker until they pass the next marker, so that steering control can be performed even if the interval between markers was unknown.



Photograph 3. Radio Wave Marker

(2) Radio wave marker sensor positioning method

With the radio wave method, an antenna inside the sensor on the vehicle transmits a 227.5 kHs radio wave towards the road and when a radio wave marker (Photo 3) receives this transmission, it returns a signal of double frequency of 455 kHz. A receiving antenna inside the sensor receives this return radio wave to detect the vehicle's position (Fig. 6).

The system detects the locations of the markers in the traveling direction by passing through the peak value. It detects the position in the lateral direction by calculating the distance between the antennas by the triangulation method based on the difference in the received strength of the return radio waves sensed by the two receiving antennas [2].

Photograph 4. Radio Wave Marker Sensor

(3) Magnetic marker sensor positioning method

Photograph 5. Magnetic Marker

With the magnetic method, magnetic markers—permanent ferrite magnets (Photo 5)—under the road surface produce magnetic fields and a magnetic sensor that detects magnetic field density installed on the vehicle determines the position of the vehicle based on the strength of each magnetic field.

The magnetic field distribution of the magnetic marker shows the unimodality that treats the center of the marker as the maximum magnetic field strength. The magnetic flux densities of the vertical component (Bz) and the vehicle width direction component (Bx) of this magnetic field are detected by the magnetic sensor and the location of the marker is detected by calculating the distance between the sensor and the marker based on the previously stipulated Bx/Bz relational equation (Fig. 4) [3).

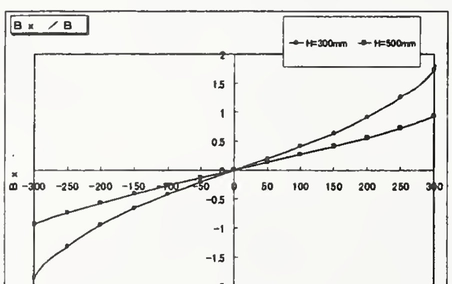


Figure 4. Bx/Bz Relationship

3.4 GPS/GIS position detection method

A system based on GPS/GIS positioning detects the position of the vehicle to control its operation by comparing rotary snow remover position coordinate data received from a GPS satellite with position coordinates of linear data for the road provided in advance by road GIS.

(1) Road GIS

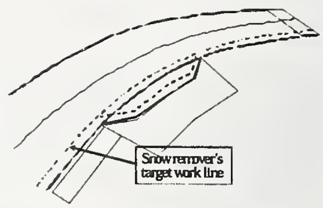


Figure 5. Road GIS Data

Road GIS is a data base of information concerning structures and traffic management used for road maintenance. To develop this system, rotary snow remover target course information was created so that it can be handled as GIS data.

Operators of rotary snow removers remove snow to widen the bare road by driving their vehicles guided by the snow removal edge (Fig. 5) Because the snow removal edge is often the curb of the sidewalk along the sides of the road, the GIS data that was used is curb data.

(2) RTK-GPS

RTK-GPS, a system capable of high-precision positioning (error radius between 2 and 3 cm), consists of two GPS receiving antennae, one a fixed station and one a mobile station (on the snow remover). The fixed station transmits correction information based on GPS data it has received to the mobile station.

The system can perform almost real time positioning, because the GPS data is corrected at a rate of 20Hz.

4. CORROBORATIVE TESTING

Corroborative testing planned to simulate snow removal on a real road was carried out using a rotary snow remover equipped with the steering control system in Hokkaido in January and February 2002.

The corroborative testing was performed in order to evaluate the applicability and practicality of each automatic steering method under the behavior characteristic of a snow remover on an actual road (low speed, lateral sliding).

4.1 Outline of the corroborative testing

(1) The rotary snow remover

The body of the rotary snow remover was the medium size model (2.2 m wide) used most frequently on national highways, and the body constructed according to normal specifications was partially modified (Fig. 6).

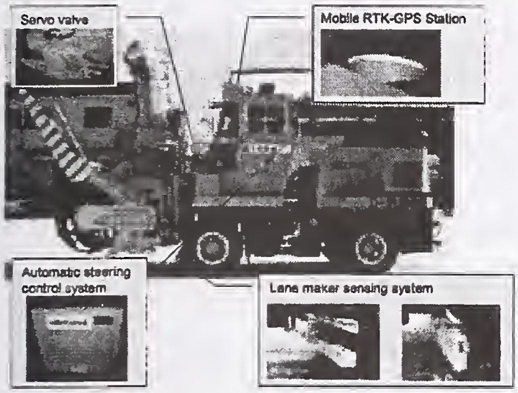


Fig. 6 Rotary Snow Remover Equipped with an Automatic Steering Support Function

(2) Corroborative test course
The corroborative test course consisted of straight sections, curves, (R 30 m), and intersections (R 12 m) so that it reproduced conditions on national highways. Its lane marker intervals were 2.0 m and 1.5 m.

The rotary snow remover work conditions were identical to those of actual work, with the work done at two speeds, 4 km/h and 0.5 km/h [4]. Snow banks were prepared to provide the work load and lateral sliding that occur when a remover is used to widen the cleared part of a road.

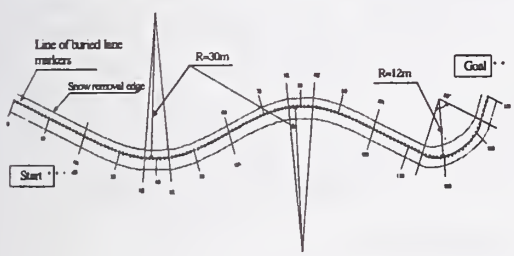


Figure 7. Layout of the Test Course

4.2 Results of the corroborative testing

The testing corroborated the lane marker sensing (radio wave and magnetic) guidance methods and the GPS/GIS positioning method.

Because the steering control system developed is a system that follows a preset target course, its control performance was evaluated as the quantity of deviation of the snow remover from its target course (quantity of lateral deviation). Table 1 shows the quantities of lateral deviation from the test results organized by method.

Table 1. Lateral Discrepancy for Each System (Example)

Stand-alone operation (working speed 4 km/h, quantity of deviation of front sensor)

Unit m

Road alignment		Straight section	Curve to the left	Curve to the right	Straight section	Inter-section
Radio wave type	Max. value	0.25	0.41	0.28	0.39	0.76
	Min. value	–0.230	– .490	– .38	0.02	0.26
	Average value	–0.05	0.06	–0.13	0.16	0.51
	Standard differential	0.12	0.23	0.19	0.12	0.16
Magnetic type	Max. value	0.06	0.35	0.37	0.42	0.91
	Min. value	–0.180	– .370	– .53	0.11	0.29
	Average value	–0.04	0.08	–0.25	0.27	0.61
	Standard differential	0.06	0.21	0.24	0.10	0.22
GPS	Max. value	0.04	0.04	0.25	0.24	0.17
	Min. value	–0.030	– .320	– .26	0.00	–0.24
	Average value	0.00	–0.080	– .04	0.07	–0.02
S	tandard differential	0.02	0.11	0.10	0.08	0.09

This table shows that overall, the quantity of deviation and the standard deviation were both small on the straight sections, but both were higher on curves and at intersections. This difference is presumably a result of the fact that in sections with a small radius of curvature, a larger quantity of steering is performed and the snow removal load is more likely to cause lateral sliding.

By method, the GPS/GIS method is superior to the lane marker method. Because the former method achieves almost real time positioning and steering control of almost 20Hz, it can respond instantly to lateral sliding to reduce the quantity of deviation. With the latter, steering is delayed because the system cannot detect lateral sliding where there are no signals between lane markers, resulting in a large quantity of lateral deviation. Therefore, testing was done by, in addition to correcting lateral sliding with the sensor, providing the coordinates of the next marker (course information) as the sensors passed the markers.

The resulting lateral sliding distribution reveals that as shown in Fig. 8, ±2 σ (σ: standard differential) at this time was ± 57 cm without course information, and it was ± 31 cm with course information: a big improvement that brings its precision extremely close to that of the GPS/GIS method.

No gap in performance between the speeds 4 km/h and 0.5/h was observed.

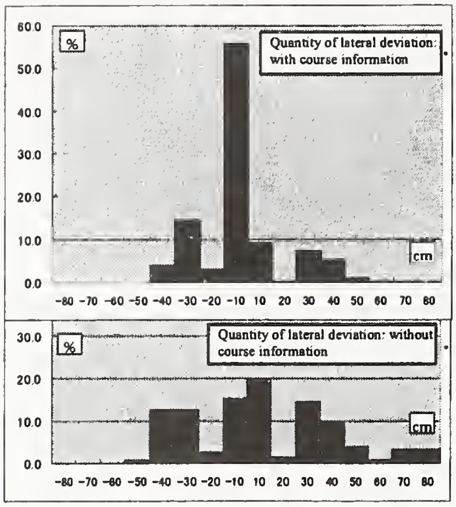


Figure 8. Lateral Deviation With and Without Course Information (Radio Wave Method)

5. CONCLUSIONS AND FUTURE CHALLENGES

It has been confirmed that automatic steering systems using lane markers or GPS, GIS, etc. can achieve a generally practical level of control of normal snow removal work by rotary snow removers. But in order to establish practical working systems, the following challenges must be resolved.

(1) Testing of the GPS/GIS method showed that its detection precision temporarily fell even in open spaces. It is assumed that this happened when the GPS satellites were concentrated in a narrow range. And during testing of the radio wave marker method, unpredictable behavior thought to be an effect of the high output transmitter on the vehicle was observed. It is necessary to provide a function that automatically stops operation without the intervention of the operator when such abnormal operating signals are produced.

(2) And as in the GPS/GIS case, it is necessary to improve the precision of control by the lane marker method by entering the lane marker coordinates to the snow remover system in advance to provide it with course information. It is also essential to record passage over each marker each time the system passes over the markers.

(3) And to further improve control precision (reduce the quantity of deviation), it is necessary to strive to quantify disturbance such as lateral sliding and the snow removal load etc. of the snow remover and to perform simulations after increasing the degree of detail of the model of snow removal work by a rotary snow remover.

6. CONCLUSION

In recent years, various organizations have attempted to apply information technology (IT) to improve construction machinery. As seen in those cases, the information handled in this case cannot be applied to other systems unless its use is premised on standardization. Because the rotary snow remover developed by this project requires new road infrastructure such as lane markers, road GIS data, and communication systems and its communication standards must be standardized, its development, has been harmonized with communication technologies in the ITS field. The authors hope that this research and development will contribute to introducing it as a working system.

In conclusion, the authors wish to express their deep gratitude to the Hokuriku Regional Development Bureau and the Hokkaido Regional Development Bureau of the Ministry of Land, Infrastructure, and Transport whose members helped with the corroborative testing and to everyone who helped with the fabrication of the system and the testing work.

REFERENCES

1) Sanpei M. Ito T.: Path planning and path tracking control of wheeled vehicles using nonlinear system theory—parking control using forward and backward movement, Collected Reports of the Institute of Systems, Control, and Information Engineers, Vol. 6, No. 1, pp. 37 to 47, 1993

2) Handa S., Nakazuka A., Tanji Y.: Lane marker system, Matsushita Technical Journal, pp. 45 – 49, October 2001

3) H. Suzuki: Development of Lane Marker System in AHSRA, 6th World Congress on Intelligent Transport Systems, November 1999.

4) Fourth edition of the Road Snow Removal Handbook: edited by the Japan Construction Mechanization Association, pp. 131

ADAPTIVE CONTROL OF A CONSTRUCTION MANIPULATOR

by

V. Gradetsky [a], M. Rachkov [b], M. Pushkin [a]

[a] *Institute for Problems in Mechanics, Russian Academy of Sciences, Vernadsky ave. 101 Moscow 117526, Russia, Phone: +7-095-4344149, gradet@ipmnet.ru*

[b] *Moscow State Industrial University, ul. Avtozavodskaya 16, Moscow, 109280, Russia, Phone: +7-095-3684883, rachk-v@mail.msiu.ru*

ABSTRACT: The paper describes an advanced control technique that uses a self-adjusting controller. The technique allows changing control parameters of the manipulator according to different technological forces and disturbances acting to the end-effector of the manipulator while performing surface treatment construction operations such as cleaning and polishing. The system is asymptotically stable for the phase error and can be implemented on a base of existed industrial robots.

KEYWORDS: adaptive control; construction manipulator, parametrical error; reference model; self-adjusting controller.

1. INTRODUCTION

Automation of some surface treatment construction operations such as cleaning and polishing by means of manipulators demands adaptive control that allows keeping required technological forces of a tool at any disturbances from surface unevenness.

There is a program that allow for efficient management of relevant operational parameters related to surface finish quality [1]. This program is intended for on-board computers in graders or pavers.

One of approaches in manipulator control is based on a calculation method by incomplete information about external parameters [2]. However, the known algorithm of control is characterized by complexity of calculations and is usually used for manipulators with two or three degrees of freedom.

Optimal positioning of a pneumatic manipulator with minimum control energy consumption is solved in [3]. This system can place a tool with high accuracy taking into account information about current position of the manipulator but some disturbances can act to the tool from treated surface unevenness while operating.

The proposed technique solves a task of qualitative control of the manipulator tool with a reference model where a wide range of possible dynamic properties can be changed. The technique allows changing control parameters of the manipulator according to different technological forces and disturbances acting to the end-effector of the manipulator while performing surface treatment construction operations such as cleaning and polishing by means of a self-adjusting manipulator controller.

2. SYSTEM WITH A REFERENCE MODEL

Let's consider the approach based on the self-adjusting system with a reference model on the basis of information about object parameters that is received while functioning. The active change of regulator parameters should be carried out on a base of this information.

The self-adjusting system with the reference model can be described by the block diagram presented in Figure 1.

The system motion is described by the linear n-order differential equation with constant beforehand-unknown factors, which can vary in a wide range scale

$$x^{(n)}(t) + \sum_{i=1}^{n}[a_i + \Delta a_i + \delta a_i(t)]x^{(n-i)}(t) =$$
$$= f(t),$$
(1)

where Δa_i - constant beforehand-unknown factors, $\delta a_i(t)$ - factors, generated by the self-adjusting circuit.

The reference model is described by the equation:

$$y^{(n)}(t) + \sum_{i=1}^{n} a_i y^{(n-i)}(t) = f(t),$$
(2)

where $y(t)$ - output signal of the reference model, $f(t)$ - control signal. It is supposed that the reference model is asymptotically steady.

The purpose the system with the reference model is to design such a self-adjustment circuit, at which $x(t) \to y(t),\ t \to \infty$ by generated $\delta a_i(t)$. The equations of the object and the reference model in the matrix form are

$$\dot{x}(t) = Ax(t) + [\Delta A + \delta A(t,x,y)]x(t) +$$
$$+ f(t), \quad x(t_0) = x_0,$$
(2)
$$\dot{y}(t) = Ay(t) + f(t), \quad y(t_0) = y_0, \qquad (3)$$

where $x(t),\ y(t) \in R^n$ - phase coordinate vectors of the object and the reference model, A - a real constant matrix (Gurvitz matrix) that can be written as

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \multicolumn{5}{c}{\dots\dots\dots\dots} \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix},$$
(4)

ΔA - real constant matrix with beforehand-unknown factors depending on the control object, $\delta A(t,x,y)$ - matrix of parameters changed by the self-adjustment circuit. The matrixes ΔA and δA can be expressed as

$$\Delta A = \begin{bmatrix} 0 & \dots & 0 \\ \multicolumn{3}{c}{\dots\dots\dots} \\ 0 & \dots & 0 \\ \Delta a_1 & \dots & \Delta a_n \end{bmatrix},$$

(5)

$$\delta A(t,x,y) = \begin{bmatrix} 0 & \dots & 0 \\ \multicolumn{3}{c}{\dots\dots\dots} \\ 0 & \dots & 0 \\ \delta a_1 & \dots & \delta a_n \end{bmatrix}.$$

The control vector is $f^T(t) = (0,\dots,0, f_n(t))$, where the top index hereinafter means a transpose operation. For an error vector $\varepsilon(t) = x(t) - y(t)$, subtracting equations (3) from (2), we have

$$\dot{\varepsilon}(t) = A\varepsilon(t) + [\Delta A + \delta A(t,x,y)]x(t),$$
$$\varepsilon(t_0) = \varepsilon_0$$
(6)

or in more convenient form

$$\dot{\varepsilon}(t) = A\varepsilon(t) + X(t)\alpha(t,x,t),$$
$$\varepsilon(t_0) = \varepsilon_0$$
(7)

where $X(t)$ - matrix, at which the last row coincides with the vector $x^T(t)$, and $\alpha(t,x,y)$ - vector of a parametrical error

$$\alpha^T(t,x,y) = \\ = \left(\Delta a_1 + \delta a_1(t,x,y),\ldots,\Delta a_n + \delta a_n(t,x,y)\right)^{\cdot}$$

The Lyapunov´s second theorem is used for a synthesis of the self-adjustment circuit. Let's choose Lyapunov´s function as the square-law form

$$V(t,\varepsilon,\alpha) = \varepsilon^T \Gamma\varepsilon + \alpha^T\alpha,$$
(8)

where the positively certain matrix Γ satisfies to the matrix equation

$$A^T\Gamma + \Gamma A = -CE.$$

The equation of the self-adjustment circuit is

$$\dot{\alpha}(t) = \theta(t,\varepsilon,x,y).$$

A derivative of the V function by this equation should be not positive, i.e.

$$\frac{dV}{dt} \leq 0.$$

The derivative of the function (8) by system (7) is equal

$$\dot{V} = -C|\varepsilon|^2 + \alpha^T X^T \Gamma\varepsilon + \varepsilon^T \Gamma X\alpha + \\ + \dot{\alpha}^T\alpha + \alpha^T\dot{\alpha} = -C|\varepsilon|^2 \leq 0 \quad ,$$
(9)

if the self-adjustment circuit is described by the following equation

$$\dot{\alpha}(t) = -X^T(t)\Gamma\varepsilon(t). \qquad (10)$$

The algorithm (10) covers the majority of known algorithms of adaptation. If vectors $x(t)$, $y(t)$, $\varepsilon(t)$ are measurable, the algorithm (10) can be implemented. Thus, the system (7), (10) is stable for $\varepsilon(t)$ and $\alpha(t)$.

3. DEFINITION OF THE SELF-ADJUSTMENT CIRCUIT VECTOR

In practice, there are objects, in which the matrix of factors depends on some number of parameters. Let the object is described by the equation

$$\dot{x}(t) = A(q)x(t) + \delta Z(t) + f(t), \quad x \in R^n,$$
(11)

and the reference model is described by the equation

$$\dot{y}(t) = A(q^0)y(t) + f(t), \quad y \in R^n,$$
$$f(t) \in R^n$$
(12)

where $q \in R^m$ - the vector of parameters, from which a matrix of object is depended, $A(q^0)$ - Gurvitz matrix, $\delta Z(t)$ - vector generated by the self-adjustment circuit. The task consists in definition of structure and algorithm of change of the vector $\delta Z(t)$ to have $\varepsilon(t) = x(t) - y(t) \to 0,\ t \to \infty$. The vector of a phase error $\varepsilon(t)$ satisfies to the equation

$$\dot{\varepsilon}(t) = A(q^0)\varepsilon + \lfloor A(q) - A(q^0) \rfloor x(t) +$$
$$+ \delta Z(t)$$
(13)

Decomposing $A(q)$ in a row, we have

$$A(q) - A(q^0) = \sum_{i=1}^{m} \Delta q_i \frac{\partial A(q)}{\partial q_i} + O(|\Delta q|),$$
$$q - q^0 = \Delta q, \quad \Delta q \in R^m,$$
(14)

where values of derivatives in (13) are calculated for $q = q^0$.

Let's enter the vector $\delta Z(t)$, generated by the self-adjustment circuit, as

$$\delta Z(t) = \sum_{i=1}^{m} \delta q_i(t) B_i(t),$$
$$B_i(t) = \frac{\partial A(q)}{\partial q_i}\bigg|_{q=q^0} x(t)$$
(15)

Here $\delta q_i(t)$ - scalar functions that is determined below.

Taking into account expressions (14) and (15), equation (13) is possible to write as

$$\dot{\varepsilon}(t) = A(q^0)\varepsilon(t) +$$
$$+ \sum_{i=1}^{m} (\Delta q_i + \delta q_i) \frac{\partial A}{\partial q_i} x(t) + O(|\Delta q|)\, x(t).$$
(16)

Let's assume, that the matrix $A(q)$ depends on the parameter q linearly. Then, last member in the equation (16) is absent and it will have a form

$$\dot{\varepsilon}(t) = A(q^0)\varepsilon(t) + \sum \beta_i(t) \frac{\partial A}{\partial q_i} x(t) \;,$$
$$\beta_i(t) = \Delta q_i + \delta q_i(t).$$
(17)

The further synthesis of the self-adjustment circuit uses Lyapunov´s function. It can be chosen as

$$V(t, \varepsilon, \beta) = \varepsilon^T \Gamma \varepsilon +$$
$$\sum_{i=1}^{m} \beta_i^2(t), \; \Gamma > 0, \; \Gamma = \Gamma^T,$$
(18)

where the matrix function Γ generally looks like $\Gamma = \Gamma(t, x, y)$.

A complete derivative of this function, taking into account equations (17), is

$$\frac{dV}{dt} = \varepsilon^T (A^T(q^0)\Gamma + \Gamma A(q^0) + \dot{\Gamma})\varepsilon +$$
$$\sum \beta_i(t)(B_i^T \Gamma \varepsilon + \varepsilon^T \Gamma B_i + 2\dot{\beta}_i(t))$$
(19)

Since a matrix Γ is symmetrical, we have $B_i^T \Gamma \varepsilon = \varepsilon^T \Gamma B_i$. Let's consider the following algorithm of adjustment of parameters $\delta q_i(t)$

$$\frac{d}{dt}(\delta q_i(t)) = \dot{\beta}_i(t) = -B_i^T \Gamma \varepsilon =$$
$$= -x^T(t) \left(\frac{\partial A}{\partial q_i} \right)^T \Gamma \varepsilon$$
$$\beta_i(0) = \Delta q_i, \; \delta q_i(0) = 0$$
(20)

Let matrix $\Gamma(t, x, y)$ satisfies to conditions

$$C_1|p|^2 \le p^T \Gamma(t,x,y) \le C_2|p|^2$$
$$p^T(A^T\Gamma + \Gamma A + \dot{\Gamma}) \le -C_3|p|^2,$$
(21)

in which $A = A(q^0)$, the matrix $A(q)$ linearly depends on parameters, $A(q^0)$ - Gurvitz matrix. Then it is possible to show, that the system (17), (20) is stable for $\varepsilon(t)$ and $\beta_i(t)$, and also is asymptotically stable for the phase error. If the matrix $A(q)$ depends on parameters not linearly, the system may be not stable.

However, described self-adjustment circuit has the following property. If the control signal is restricted, the phase error is also limited, and the error can be so small as desired, if the parametrical error between the object and the reference model is small enough. It results from general stability theorems.

4. CONCLUSIONS

The advanced control technique that uses the self-adjusting controller is developed. The technique allows changing control parameters of the manipulator according to different technological forces and disturbances acting to the end-effector of the manipulator while performing surface treatment construction operations such as cleaning and polishing.

The system is asymptotically stable for the phase error in a linear case. In non-linear case, the phase error is also limited, and the error can be so small as desired, if the parametrical error between the object and the reference model is small enough.

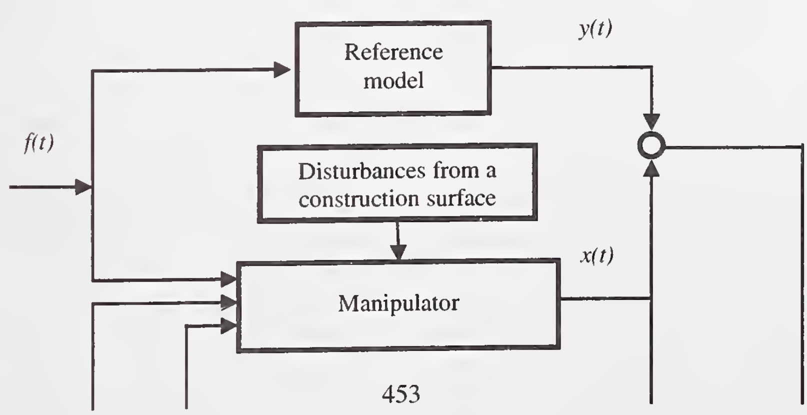
The described control technique can be implemented on a base of industrial robots RM-01 or PUMA 560 that has six rotary joints and can perform different construction operations with complex trajectories of the tool.

5. REFERENCES

1. Skibniewski, M., Michalowski, S., Sobczyk, A.: "Predictive control systems in construction equipment", 18th ISARC Proceedings, 2001, Krakow, Poland, pp. 277-282.

2. Fu, K.S., Gonsalez, R.C., Lee G.S.G.: "Robotics: Control, Sensing, Vision and Intelligence", McGraw-Hill Book Company, New York, 1987, 624 p.

3. Rachkov, M., Marques, L., Almeida, A.T.: "Optimal Control of Pneumatic Manipulators for Construction Application", 18th ISARC Proceedings, 2001, Krakow, Poland, pp. 253-258.



JUST-IN-TIME CONTINUOUS FLIGHT AUGER PILES USING AN INSTRUMENTED AUGER.

by

N. Mure[1], J. N. Scott[2], D. W. Seward[3], S. D. Quayle[4], C. R. I. Clayton[5], M. Rust[6].

ABSTRACT:

Continuous flight auger (CFA) piles are constructed by drilling a CFA auger into the ground and, on reaching the required depth, pumping concrete down the hollow stem as the auger is steadily withdrawn. Current practice to predict the bearing capacity of CFA piles is to estimate the undrained shear strength-depth relationship for the overall site, and use a total stress analysis to predict a general pile bearing capacity. This analysis is often based on sparse site data collected from a location remote from the pile. This paper investigates the exploitation of new technology to enable a new improved approach to the procurement, design and validation of (CFA) bored piles. The ultimate target is that the final length of the piles will be determined on site, as they are constructed, and will be optimised to suit the actual local ground conditions.

KEYWORDS: data logging; instrumentation; pile foundations; process optimisation.

1.0 INTRODUCTION:

CFA piles are formed by shearing the soil with a continuous 'corkscrew' like auger, and then replacing the soil with concrete as the auger is withdrawn in a controlled fashion. Finally reinforcement is inserted. They offer a highly efficient and cost-effective solution under certain conditions, however there remains some variation in practice between piling contractors, and in certain situations there is a need for improved confidence in final integrity of the pile. Increased understanding of the data collected on-board during the actual placing of the pile could provide this increased confidence and reduce risk, it would open the way to a radical reform of the business process, with significant changes to the relationships between the bodies involved, and the prospect of economic benefits to all parties [1]. The ultimate target is that the final length of the piles will be determined on site, as they are constructed, and will be optimised to suit the actual local ground conditions.

2.0 CFA PILES:

At face value the installation of a CFA pile seems simple, a hollow stemmed auger is drilled into the soil and concrete is then pumped into the ground through the stem as the auger is slowly withdrawn. However, the action of drilling the pile is a complex one to

[1] Stent Foundations Limited, Birchwood Way, Cotes Park West, Somercotes, Derbyshire, DE55 4PY, U.K., Tel: +44 1256 763161, Email: norman.mure@stent.co.uk

[2] Stent Foundations Limited, Osborn Way, Hook, Hampshire, RG27 9HX, U.K., Tel: +44 1256 763161, Email: Jason.Scott@stent.co.uk

[3] Engineering Department, Lancaster University, Lancaster, LA1 4YR, U.K., Tel: +44 1524 593010, Email: d.seward@lancaster.ac.uk

[4] Engineering Department, Lancaster University, Lancaster, LA1 4YR, U.K., Tel: +44 1524 593095, Email: s.quayle@lancaster.ac.uk

[5] Department of Civil and Environmental Engineering, University of Southampton, Highfield, Southampton, SO17 1BJ, U.K., Tel: +44 23 8059 2841, Email: cric@soton.ac.uk

[6] Department of Civil and Environmental Engineering, University of Southampton, Highfield, Southampton, SO17 1BJ, U.K., Tel: +44 23 8059 3492, Email: m.rust@soton.ac.uk

model and understand. It is not always obvious what affect the continuous flight auger has on the ground.

With CFA, the process remains unseen and there is no way of carrying out visual inspections of the bore. We are therefore forced to draw conclusions from past experiences, as to the likely consequences of installing certain piles in certain conditions.

The process of installing a CFA pile comprises of three phases or parts. First there is the boring of the pile, then concrete is pumped into the system pre-charging the lines and auger and finally the pile is concreted. The nature of the rig, the ground conditions and the augers all have an impact on each of these phases.

2.1 Boring the Pile

As the auger is rotated into the ground material from the auger tip is loaded onto the flights. The auger is carrying out three distinct actions - cutting, displacement and transportation of soil.

Figure 1 demonstrates the auger being screwed into the ground. As the auger moves downwards the surrounding soil has to be displaced by the volume of the auger. This is typically 15-20% of the volume of the bore.

As the ground is excavated at the cutting head it bulks and a ribbon of spoil moves up the auger. Because of this bulking there is always sufficient material available to fill the flights. The full flights now support the bore, over the length of the auger.

It would be ideal if CFA piles could be completely installed in this way, with the auger being continually screwed into the ground. However, even the largest rigs do not have sufficient power to overcome the frictional forces that are generated by displacing the auger volume into the surrounding soil.

Instead, it is necessary to restrain the auger during the boring process. This allows some soil to travel up the flights (transportation) with no further material being loaded onto the auger at the bottom. If the flights were allowed to empty in this way then the auger could start to load from the sides of the pile bore, a phenomenon known as side loading.

Side loading would loosen the surrounding soils which is detrimental to the skin friction of the finished pile. In fact many CFA piles have been observed to have lower skin friction than expected from calculation. In addition on some sites skin friction has been observed to vary from one pile to another by as much as 100%, which illustrates just how critical the boring of the pile is to its performance.

2.2 Pre-charging

This refers to the practice of pumping sufficient concrete into the auger stem in order to fill it before extraction of the auger begins. Typically the required pre-charge volume is between 150 and 300 litres.

After boring and while pre-charging it is important that the auger is not excessively rotated. In granular soils this action can have a detrimental affect not only on the shaft friction because of side loading, but also on the end bearing capacity of the pile due to the material at the base being loosened.

Prior to pre-charging it is common practice to lift the auger slightly off the bottom of the bore in order to allow the bung or "clack" at the auger tip to open (to allow the concrete to flow through). Experience shows that this lift should be kept to a minimum of 100mm.

2.3 Concreting

Once the auger has been pre-charged and the bung is blown, extraction can begin. It is good practice to rotate the auger during the initial stages of concreting in order to carry concrete and debris up onto the auger. Once this has been done and concrete pressure is observed at the swan-neck then the auger can be withdrawn, balancing the rate of extraction with the amount of concrete pumped through it.

The concrete supply pressure is usually measured at the top of the auger stem at the swan neck. This means that for pressure to be measured the auger string has to be full of concrete. With a stem of 25m in length this translates to a pressure of about 500kN/m2 or 5 atmospheres at the auger tip.

Given that measuring pressure is not always possible, concrete flow rate versus auger extraction becomes the best guide to producing sound piles. Over-break is that concrete used over and above the net volume of the bore. Oversupply should always be positive including towards the pile top where the ground is loosest, to ensure that concrete always flows up the flights. Otherwise contamination is possible, particularly in wet soils.

Over-break targets are critical for the correct construction of a CFA pile. In stiff soils pulling to high targets, say of 25% or more, may be impossible as the bore cannot yield to take the oversupply. In this case concrete will eventually escape to the ground surface or, more likely, the high pressure will lead to a blockage in the auger. Conversely, low targets in loose or soft ground could lead to defective piles, as the concrete will slump and leave the bottom of the auger tip exposed. It is therefore important that the over-break target is set for each contract and modified throughout the job as necessary.

3.0 STENT INTEGRATED RIG INSTRUMENTATION SYSTEM (SIRIS):

SIRIS is first and foremost a system for monitoring the construction of continuous-flight-auger (CFA) piles. SIRIS has been developed entirely by Stent and has been in service since about 1998. It is currently being used on all rigs in the CFA fleet [2].

The purpose of SIRIS is to assist the rig driver with constructing piles by providing him with a detailed picture of various parameters during the construction process. These parameters include auger depth, rotation, concrete pressure, concrete flow and over-supply.

In addition, SIRIS produces a complete record of the construction process for each pile, including the occurrence and cause of any delays. This record or "pile log" is then used for project management purposes and to produce a graphical record of each pile for presentation to the client. These client plots can be provided on paper or in the form of an electronic record along with the software necessary to view them.

SIRIS is effectively made up of three parts:

1. A computer on each of the CFA piling rigs and a number of sensors.
2. A database for storing, reviewing and managing pile logs.
3. A viewer for use by engineers or clients to review pile logs.

3.1 Rig Computer

Each piling rig is fitted with an industrial IBM compatible PC running the Windows NT operating system. This PC is connected to a full-colour, touch-sensitive high-brightness screen through which the driver controls the whole system. The screen is bright enough to be used in direct sunlight and the system is controlled via a user-friendly system of coloured buttons.

This PC is also connected to a number of sensors distributed about the piling rig as shown in Figure 2.

3.1.1 Concrete Pressure

A pressure sensor at the top of the auger, attached to what is referred to as the swan-neck, is used to measure the pressure of concrete pumped through the auger during the concreting phase of piling.

3.1.2 Depth

Two sensors on the rotary table or at the cat-head (the top of the mast) are used to measure vertical movement of the rotary table in order to calculate auger depth.

3.1.3 Rotation

Another two sensors on the rotary table are used to measure the number of revolutions of the auger and the direction in which it is rotating.

3.1.4 Concrete Flow

This is measured by a sensor in the line that detects the strokes of the concrete pump and therefore allows the volume of concrete to be calculated.

3.1.5 Torque

The pressure of the hydraulic fluid delivered to the rotary table is measured in order to give a rough indication of the torque being delivered by the auger during boring.

During the piling process SIRIS interrogates these sensors fifty times per second and uses the data to present the driver with a clear picture of what is happening. The sensor values are also used to create a detailed log of the whole process. Information is recorded for every 0.1 metres of the pile during both boring and extraction. In addition delays are also recorded, along with the cause of the delay as indicated by the driver from a pre-defined list.

3.2 The SIRIS Database

The pile logs recorded on the CFA rig computers are downloaded on a daily basis, either by floppy disk or (increasingly often) by GSM modem. They are then copied onto the SIRIS database which resides on a server on the Stent internal network

As well as providing a mechanism for storing, retrieving and reviewing pile logs the database generates a number of statistics which are used to assist the management of piling contracts - such as rates of production and a breakdown of the cause of delays. It also generates statistics on a historical basis which can then be used to improve the estimates for future jobs. It is because of this use of data for other internal processes that SIRIS is referred to as an integrated rig instrumentation system.

3.3 Pile Log Viewers

As well as the SIRIS database which is only accessible on Stent's network, there are also two other "stand alone" software packages that can be used for reviewing pile logs. One of these is used by Stent's engineers when they do not have access to the network. The other is a simplified version of this software that is intended for use by clients for reviewing and printing pile logs that have been passed to them electronically.

4.0 JUST-IN-TIME PILING:

In his 1984 Rankine lecture Wroth discussed different in-situ tests in which the parameter Undrained Shear Strength (τ_u) is measured. He showed how the tests were fundamentally different and compared them to the unconsolidated undrained triaxial test. Differences such as the directions and freedom of rotation of the three principal stresses were highlighted. He described the different test results as a hierarchy of τ_u. He concluded, "it is imperative for a designer to recognise this hierarchy, and to select a strength which is appropriate to the analysis or the design procedure being used" [3]. For the specific case of Continuous Flight Auger (CFA) piles the industry relies mostly on the unconsolidated undrained triaxial and/or standard penetration test to obtain a general view of the τ_u with depth over a site. This is then used to predict the performance of all the piles on a site using an undrained analysis. The problems associated with this process are well known in the field of geotechnical engineering and will not be discussed in this paper. The results of an Imperial college investigation showing the inadequacy of current pile design practices are not surprising given the current design methods. Sixteen entries from seven countries were received for the prediction of load capacity and displacement of a jet grouted and a control pile. The predictions for the load bearing capacity of both piles differed by almost an order of magnitude [4].

4.1 "Smart" Auger

Sensors on the piling rig vehicle could obtain much of the information required for JIT-pile, such as auger torque, and this is the obvious approach. However there are significant problems:

- Large piling rigs tend to be custom manufactured and vary in details such as hydraulic pump performance. This means that each individual rig would require independent setting-up and calibration.
- Time taken for setting-up, calibration and maintenance of the sensors would reduce the availability of a high-cost asset.
- The use of variable displacement hydraulic pumps and motors means that the relationship between system pressure and torque may not be that straightforward.
- Moving JIT-pile equipment from one rig to another becomes a major task.

The solution is to concentrate as much sensing as possible on the auger itself. Augers contain a short link that can be engineered into a focus for data collection and communication. A link has been prepared by the addition of strain gauges to measure both torque and axial force in the auger. Calibration has been carried out in the laboratory at Lancaster (Figure 3). The information is then relayed to the rig instrumentation system by radio link telemetry. This is the so-called "Smart Auger" concept and represents an important commercial spin-off from the research. Such systems can be easily moved between rigs and serviced independently of the expensive plant.

The adding of additional instrumentation and sensors to the auger itself allows direct measurement of torque and vertical load information during the boring stage of pile production. These readings can then be interpreted and used to estimate the undrained shear strength (τ_u) of the surrounding soil.

4.2 Method

The auger is advanced one pitch every revolution or "corkscrewed" into the soil for a certain distance, to ensure the least possible disturbance to the soil. The auger can do this under its own weight. The winch is then applied to pull back on the auger stopping penetration, while the auger is allowed to rotate at a specified rate. The cylinder of soil penetrated by the auger is sheared. The torque as well as the pullback force on the auger is measured and the shear stresses imposed by these on the cylinder of soil being sheared is calculated. The resultant maximum shear stress is then assumed to be related to the undrained shear strength of the soil. Model tests at Southampton University have been used to test this theory.

4.3 Results

Work is continuing to validate the lab based model tests with the production of a full sized auger section equipped with the necessary sensors and instrumentation. A test rig capable of applying the large forces has been constructed to permit full calibration and testing of the auger section with field trials due to commence shortly.

5.0 CONCLUSIONS:

If the information gathered proves not to be accurate enough for design purposes it still provides a useful amount of information for validating every pile on site and picking up anomalies in subsurface conditions. Furthermore it could lead to the automation of the drilling process, eliminating the influence of the operator on performance of the pile.

To date, all the indications are that an appropriate relationship between drilling resistance and pile strength has been established. Current testing work is underway at Southampton to verify the results and the work will shortly progress to site testing and verification.

The effective exploitation of 'difficult ground' will increase in importance as old sites are redeveloped for new uses. This means that the demand for piled foundations will increase, and it is therefore timely to address the issue of how new technology can improve the process.

6.0 ACKNOWLEDGEMENTS:

Throughout the grant period Stent Foundations Limited, under their own resource, developed their own rig instrumentation system known as SIRIS, and it was decided that both the collection of additional data and the final JIT-pile system should be compatible with this. The JIT-Pile funding has been through the UK's EPSRC - grant GR/M90450.

7.0 REFERENCES:

1. D. W. Seward, S.D. Quayle, N. Mure, J. N. Scott "*On Site Pile Instrumentation and its Integration into the Management Process*", 17th IAARC/IFAC/IEEE Int. Symp. on Automation and Robotics in Construction, Tokyo, Japan, (2000)

2. J.N.Scott "*Integrating Piling Rig Instrumentation with the Site and Project Management Process*", International Symposium of Tunnel Construction & Piling, 1999.

3. Wroth, CP. "*The interpretation of in situ soil tests*." Geotechnique; 1984. Vol 34, No 4 pp 449-489.

4. New Civil Engineer. "*Scattering predictions*" Dec 1999 pp 34.

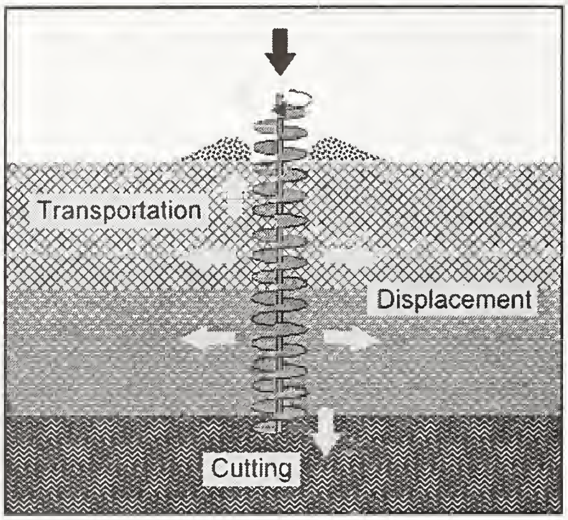


Figure 1 – CFA processes during the boring phase.

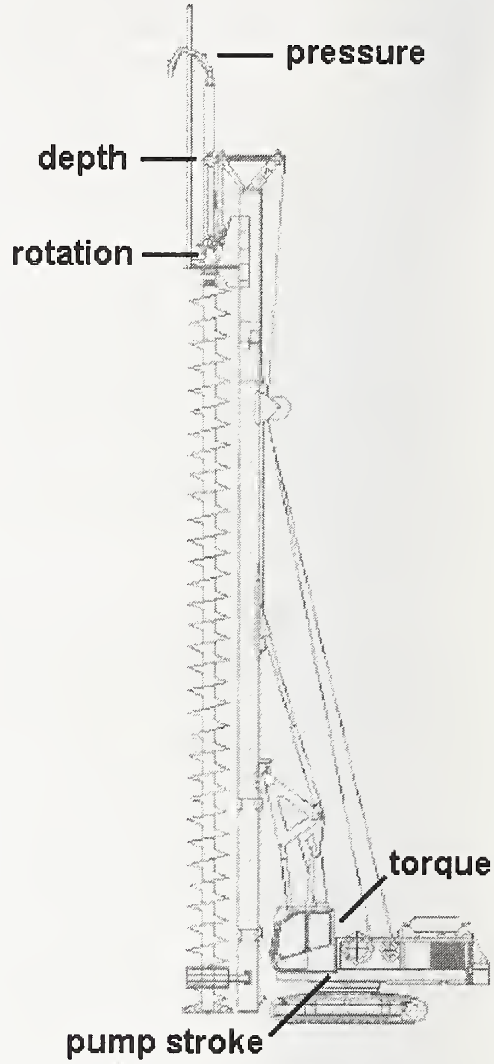


Figure 2 – CFA rig based sensors

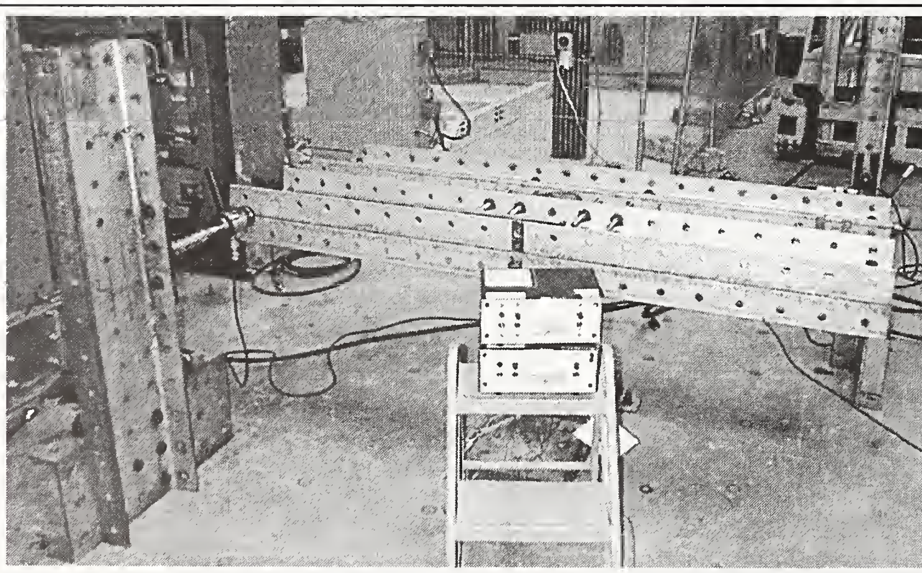


Figure 3 – Auger instrumentation testing at Lancaster University.

GPS-BASED WIRELESS COLLISION DETECTION OF CONSTRUCTION EQUIPMENT

Amr A. Oloufa[1] , Masaaki Ikeda[2] & Hiroshi Oda[3]

1- Associate Professor & Director of Construction Engineering
Department of Civil & Environmental Engineering, University of Central Florida.
P.O. Box 162450, Orlando, FL 32816-2450, USA.
Email: aoloufa@mail.ucf.edu

2,3- Fujita Technical Center, Fujita Corporation
Ono 2025-1, Atugi, Kanagawa pref., Japan
Email: m-ikeda@fujita.co.jp [2] Oda@fujita.co.jp[3]

Abstract: This paper reports on research related to avoiding collisions in construction sites using differential GPS technology. In this project, the researchers developed and implemented a system where GPS technology was used in tracking a single vehicle and relaying its information to a central server. Using another simulated vehicle, the server evaluated collision scenarios and sent cautionary messages to the roving vehicle if a collision is impending. The paper concludes with a summary of the application, along with a discussion of the limitations of GPS technology and the required augmentation by other technologies.

Keywords: GPS, collision detection, construction, robotics, automation

INTRODUCTION

Construction sites with regular and robotic equipment suffer from the potential of collisions between the various types of equipment. Such collisions are extremely costly in terms of their potential injury and other related costs and wasted time. The situation is further aggravated in sites with a large number of robotic equipment that is remotely operated from relatively long distances with the aid of cameras. In this situation, construction robotic equipment operators are often limited in their viewing fields due to insufficient cameras or bandwidth and other transmission challenges. The aim of this project is to detect and prevent impending collisions through the application of the Global Positing System (GPS) and wireless communications.

COLLISION DETECTION TECHNOLOGY

Several technologies exist for collision detection and avoidance. They differ in their cost, size, response time, reliability, and effective operational range. Ultrasonic technologies rely on high frequency devices. They have a low implementation cost, small size, and have a fast response time. They are not reliable under some conditions, and their range is only a few feet. Infrared technologies have been used for a while. They have a very small implementation cost and are small in size. However, they exhibit a high response time, are not very reliable, and their range is also very short. Radar technologies are perhaps the most effective for collision detection. Improvements have led to size reductions and high reliability. However, their cost is relatively high. Vision technologies have also been used for collision detection. However, their cost is extremely high due to heavy computational requirements. They also suffer from low reliability in some lighting conditions.

GPS technologies offer a multitude of benefits and their costs have been continuously decreasing. The major benefits of using GPS are that these technologies are not dependent on line-of-sight issues (to other vehicles), which is one of the major limitations of all the technologies listed above.

GLOBAL POSITIONING SYSTEM

Global Positioning System (GPS) is a satellite based radio-navigation system. There are 24 GPS satellites orbiting the Earth and transmitting radio signals. Based on measurements of the amount of time that the radio signals travel from a satellite to a receiver, GPS receivers calculate the distance and determine the locations in terms of longitude, latitude, and altitude, with great accuracy. GPS was created, and is controlled by the U.S. Department of Defense (DOD) for military purpose, but is available to civilian users worldwide free of charge.

GPS can be used in various areas such as air, land, and sea navigation, mapping, surveying and other applications where precise positioning is required. The system inherently has no limitation in speed or altitude, but U.S. DOD requires that commercial receivers be limited to operate below about 900 knots and 60,000 feet (18,000 meters).

Accuracy of GPS is the degree of conformance between the estimated or measured position, time, and/or velocity of a GPS receiver and its true time, position, and/or velocity as compared with a constant standard. Radio navigation system accuracy is usually presented as a statistical measure of system error and is characterized as predicable, repeatable, and relative accuracy (*Fundamentals of GPS*, 1996).

The accuracy of GPS receiver is affected by errors caused by natural phenomena, mechanical failure of elements in the system, or intentional disturbance.

For the purpose of collision detection, real-time high position accuracies (sub-meter) are required.

TELE-EARTHWORK SYSTEM

In 1994, Mount Fugen volcano, situated in southern Japan, erupted (see Figure 1). Lava flows from the volcano flowed downhill threatening the town of Shimabara. A project was developed to construct two canals to channel away future flows into the Sea of Japan.

Since work of this nature is carried under the constant threat of lava flows, it was desirable to use in construction, an automated tele-earthwork system, remotely controlled from a safe distance (See Figure 2).

The Fujita Corporation, a large Japanese Construction conglomerate, developed and implemented a Tele-earthwork system. This system consists of dozers, backhoes, trucks, and other ancillary vehicles and equipment. The system is designed to be remotely operated and has been used in several projects as shown in Figure 3.

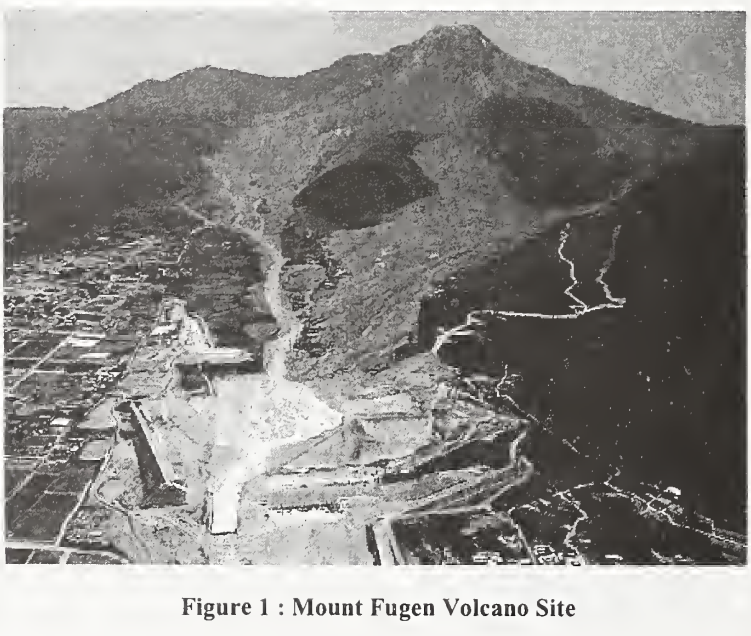


Figure 1 : Mount Fugen Volcano Site

1994: 1st trial at Mt. Fugen
(Excavating and transporting earth and sand) 6,500m3

1994□ 2nd trial at Mt. Fugen
(Excavating and transporting earth and sand) 16,000m3

1994-1995□ 3rd trial at Mt. Fugen
(Excavating and transporting earth and sand) 100,000m3

1997: Recovery from disaster of a slope collapse at Kumamoto

1997: Recovery from disaster of a slope collapse at Nagano

1997: 1st trial of unmanned constructing dam at Mt. Fugen

1998: Recovery from disaster of a slope collapse at Akita

1998: 2nd trial of unmanned constructing dam at Mt. Fugen

1999: 2nd trial of unmanned constructing dam at Mt. Fugen

COLLISION-RELATED ISSUES

In automated construction applications, especially in the case of remote operation, the lack of true visual and depth perception increases the likelihood of collisions between equipment involved in the operation.

And while the literature is full of work related to optimum path selection for robotic and automated equipment, the implementation of these algorithms, especially in earthmoving sites with their constantly changing topology and large number of equipment, have proven to be impractical.

The mostly widely used technology for collision detection is radar-based. And while this technology is excellent for avoiding outright collisions, it is ill-suited for applications that involve work in dirty environments, with large objects or terrain obscuring other vehicles that are not detected by radar. Here GPS technology comes to the rescue.

Because the technology does not rely on line of sight to other monitored vehicles, the researchers developed an architecture for sensing and warning vehicles of impending collisions as shown in Figure 4.

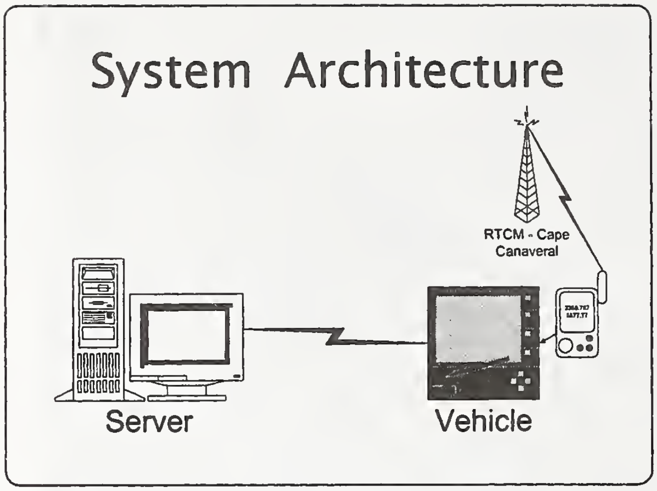


Figure 5 shows the schematic of an advanced technology construction system developed by Fujita Corp. All of the construction equipment at the site operates without on-board drivers. The vibrating roller moves autonomously, aided by GPS. Another backhoe, bulldozer, and two dump trucks are remotely controlled, each by a different operator.

While operating the equipment, each remote operator had to be watching several screens showing images from a camera on the vehicle, and another remotely controlled camera at the site. This denies operators the improved capability of controlling equipment in reference to "true" 3D perspectives. To avoid collisions, this architecture required increased vigilance by operators, and frequent verbal warnings among operators.

In this project, the proposed collision detection system would have substantially improved construction efficiency, and reduced operator overload.

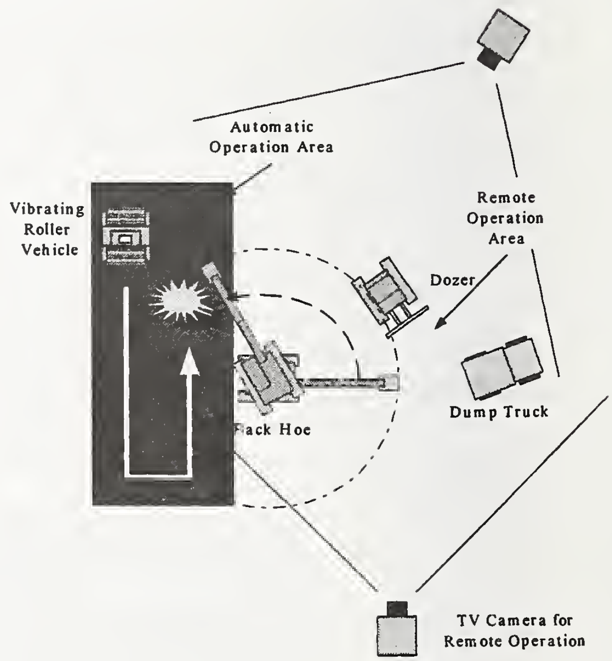


Figure 5: Arrangement of equipment in site

HARDWARE ARCHITECTURE

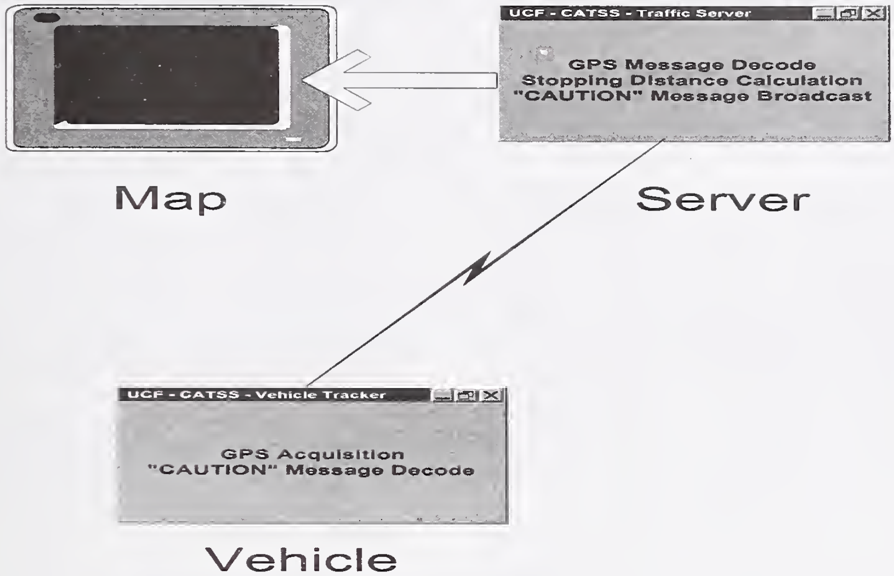
The hardware is comprised of two subsystems, a server subsystem, and a mobile (rover) subsystem. The server subsystem consists of a PC-based server connected to a radio. The rover (equipment side) consists of a differential GPS system receiver, GPS, a laptop, and a radio.

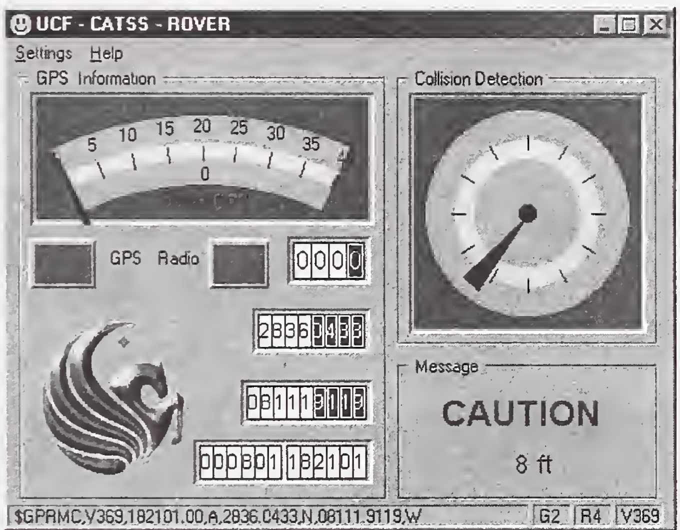
SOFTWARE ARCHITECTURE

The software architecture is shown in Figure 6.

The software consists of:

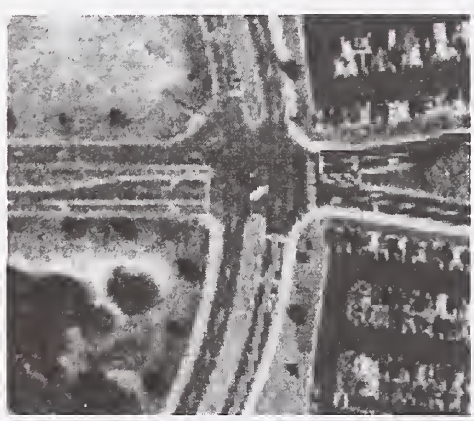
1) Vehicle Software
This software receives the DGPS message and then sends it to the server. The position is formatted according to the RMC sentence.





2) Server Software

This software receives position information from the construction equipment, and calculates potential collision scenarios with the simulated vehicle. If a potential collision is detected, the equipment operator is alerted with the speed and direction of the simulated vehicle (see Figure 7). The operator can also view construction equipment position, and track on a an aerial view of the construction site, as shown in Figure 8.



vectors representing two moving vehicles. Each vector is defined by a point And a direction. In this case, The GPS position of the vehicle (i.e. vehicle location), and the vehicle bearing (also from GPS input) define each vector.

After the intersection point is computed, and knowing the vehicles' speeds from GPS, the program calculates the distance from the potential collision point to each vehicle location. The program also calculates the braking distance required for each vehicle in its operational scenario. If the braking distance required approaches the distances above (within a specified tolerance value), the server then issues potential collision alerts to the vehicles in question transmitting an alert message, along with the direction and distance of the vehicle in question.

CONCLUSION

This paper presents research aimed at utilizing GPS technology and wireless communications for collision detection on construction websites. The technology can be applied in both automated as well as traditional construction sites.

The technology presented here seems to have a lot of promise, however, several areas such as optimum system architecture, signal reliability, GPS accuracy, and potential differential signal latency and communications issues have to be evaluated.

ACKNOWLEDGMENTS

The authors wish to thank the Center for Advanced Transportation Simulation Systems (CATSS) of the University of Central Florida, and the Fujita Corporation for sponsoring this research.

REFERENCES

Dana, P. H. *Global Positioning System Overview* TX, 1995.

Do, W. S., A. Oloufa and H. Thomas. *Evaluation of GPS devices for a Quality Control System for Compaction Operations.* Transportation Research Record # 1675 pp. 67-74, 1999.

Fundamentals of GPS. Arlington: VA, Navtech Seminars, Inc., 1996.

GLONASS Group, "*About GLONASS*", MIT Lincoln Laboratory, http: //satnav.atc.ll.mit. edu/glonass/about_glonass.html, December 1997.

Hurn, J. *Differential GPS Explained.* Sunnyvale: CA, Trimble Navigation, Ltd., 1993.

Kennedy, M. *The Global Positioning System and GIS: An Introduction.* Chelsea: MI, Ann Arbor Press, Inc., 1996.

Li, C. *Development of A Tracking System for Pavement Compaction*, Master of Science Thesis, Department of Civil Engineering, Penn State University, 1995

Oloufa, A. A. "*A GPS-based System Architecture for the Quality Control of Asphalt Compaction*". IEEE Robotics and Automation Magazine. March 2002. Vol. 9, No.1, pages 29-35.

Van Diggelen, Frank, "*GPS and GPS+GLONASS RTK*", New Product Descriptions Session, ION-GPS 97, Kansas City, MO, 16-19 September 1997.

Y.F. Ho, H. Masuda, H. Oda, L. W. Stark "*Distributed Control for Tele-Operations*". Advanced Intelligent Mechatronics. 1999.

PROTOTYPE IMPLEMENTATION OF AN AUTOMATED STRUCTURAL STEEL TRACKING SYSTEM[1]

by

Karen M. Furlani[2], DeWitt T. Latimer IV[3], David E. Gilsinn[4] and Alan M. Lytle[5]

ABSTRACT: This paper discusses the prototype implementation of a system developed by researchers at the National Institute of Standards and Technology (NIST) for the transfer of real-time on-site metrology and metrology-based data for tracking steel frame construction. The purpose of the implementation is to demonstrate the feasibility of automatically transferring information from the construction site to project management databases and associated applications, specifically for the identification and tracking of structural steel subsystems. The successful integration and implementation of the on-site field data collection system with a project information management system enables a field worker to identify and track a steel member's final position and orientation on the job site.

KEYWORDS: construction automation, 3-D coordinate measurement systems, project information management systems, communication protocols, coordinate frame transforms, position and orientation determination, steel tracking

1.0 INTRODUCTION

Many existing technologies that aid in the automation of construction component tracking systems are limited in their use by a lack of construction industry standards supporting interoperability between various hardware and software systems. In addition, to achieve the precise positioning of components to a level of accuracy sufficient for the placement of structural steel, standard methods are needed for the registration and calibration of 3-D coordinate measurement systems and for the determination of part position and orientation. To assist the construction industry in overcoming these obstacles and achieve a fully integrated and automated environment, the National Institute of Standards and Technology (NIST) has on-going research in several related areas. Researchers in the Construction Metrology and Automation Group (CMAG) are involved in the fundamental research and development of position/orientation tracking systems, sensor interface protocols for construction data telemetry and construction site simulation. The Computer Integrated Construction (CIC) group is doing research on the visual representation and simulation of construction and building related models, activities, and processes.

CMAG and CIC have collaborated on a joint implementation to demonstrate the feasibility of transferring information from the job site to project management databases and associated applications in a real-time and automated fashion. The purpose of the combined systems is to enable a field operator to identify, register and spatially visualize the final position and orientation of select structural steel members at the job site. Prior to the implementation, the Component Tracking (Comp-TRAK) system functioned independently of the Project

---

[1] Certain commercial equipment, instruments, or materials are identified in this report in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

[2] Former member of the NIST Construction Metrology & Automation Group (CMAG). Current address: 1201 Cutter Lane, Boulder, CO 80302, (303) 440-8433, karen@furlani.com

[3] Former NIST CMAG member. Current address: 4902 Forbes Ave, Box 114, Pittsburgh, PA 15213, (412) 901-0549, dl4s@andrew.cmu.edu

[4] Optimization and Computational Geometry Group member, NIST, 100 Bureau Dr., Gaithersburg, MD 20899, (301) 975-3491, david.gilsinn@nist.gov

[5]CMAG member, NIST, 100 Bureau Dr., Gaithersburg, MD 20899, (301) 975-6078, alan.lytle@nist.gov

Information Management System (PIMS), relying on post-processing to achieve automated tracking of steel members at the construction site. Utilizing standard interface protocols to integrate the Comp-TRAK system with the PIMS enables seamless communication between the various system components.

2.0 SCOPE

In earlier work, CMAG proposed an overall system architecture [1] and developed a prototype system that experimentally demonstrated 3-D spatial tracking of discrete components in real-time and under controlled conditions to viewers at a remote location [2]. The prototype field data collection system was field tested at the Building 205 Emission Control System (ECS) Addition – a US$6 million project on the NIST, Gaithersburg campus [3].

Previous work in communications research investigated the applicability of IEEE 1278 as a protocol for exchanging information about changes to the construction site [4]. Since the IEEE 1278 standard is written for distributed interactive simulation, it does not contain all the features necessary to describe the results of building a world model from sensors. Earlier work focused on answering the data pedagogy questions, specifically, the "who, what, where, and when" criteria for the types of data to be communicated [5]. These previous projects cumulated in small prototypes that lead up to this project.

This paper discusses advances in the field data collection system, site registration method, coordinate transform routine, part locator routine and the project information management system, as well as the underlying communications that enable the various systems to function in a seamless fashion. The prototype implementation was limited to capturing the "as-built" status of structural steel members during the construction of the Tower #2 Wind Frame Assembly at the Building 205 ECS Addition.

3.0 SYSTEM ARCHITECTURE

Overviews of the system architecture and the flow of operations are provided in Figures 1 and 2, respectively. The following sections describe the function of the sub-systems.

3.1 Part Tracking Overview

Structural steel members scheduled for arrival on the Bldg. 205 ECS construction site are each tagged with a bar code at the galvanizing plant prior to shipment. Reference data for each tagged part is concurrently pre-loaded into the PIMS database. Steel member identification data are directly scanned into a rugged, wearable computer with wireless access to the PIMS database hosted on a remote server. A database query following the scan provides additional part-related information as well as a 3-D VRML (Virtual Reality Modeling Language) model of the steel member. User-friendly Web browsing software then guides the field worker through the process of measuring the position of key fiducial points with a long-range, laser-based 3-D coordinate measurement system. These fiducial points are pre-specified measurement locations defined at readily identifiable features of interest such as corners. Three fiducial points, of which two must be non-collinear, provide a sufficient set of locations to establish the position and orientation of the member in 3-D space. The measured points are transformed from the local coordinate frame to a globally referenced site coordinate system. In this implementation, the member's site frame position and orientation (pose) is calculated via a Part Locator Routine hosted remotely. The steel member's identification data and local fiducial measurements are wirelessly transmitted to this routine that updates the project database with the member's site-frame pose. The same data are transmitted back to the field worker's wearable computer to update the user's visual display.

3.2 Field Data Systems User Interface

All operations conducted on the field data collection system employ a web-based interface. This includes user I/O and remote access to the part tracking routines and PIMS. This interface format was chosen over a proprietary application because it allowed inherent cross-platform compatibility and relatively simple page creation and modification.

3.3 Project Information Management System

The Project Information Management System (PIMS) allows information relating to structural steel members to be stored, accessed and modified. There are four essential parts of the PIMS system used during the implementation: an object-oriented database, a JAVA-based database server, a Common Object Request Broker Architecture (CORBA) compliant trading service, and a database administrative tool.

The database was developed based on a minimal physical schema derived from the CIMSteel Integration Standard (CIS) CIS/2 logical schema [6]. For the implementation, part definition data based on the CIS/2 representation of steel members was pre-loaded in the PIMS database via the administrative tool. This information is used to simulate structural steel parts in the ECS Tower #2 sub-assembly in VRML. The CORBA is used as the underlying communications infrastructure, allowing users to connect to the PIMS server and get information to and from the database.

4.0 POSE DETERMINATION

4.1 3-D Coordinate Measurement System

To successfully deploy the tracking system in the field, in-situ measurements of the steel frame members must be registered to a globally referenced site coordinate system for part pose determination. The field tracking system uses the Vulcan 3-D coordinate measurement system by ArcSecond Inc. to measure the local frame position of each target fiducial point on the part. This system uses two rotating laser transmitters and a receiving wand to calculate wand tip position via triangulation. The calibration method utilized for the implementation employs a proprietary variation of standard photogrammetry bundle techniques to calculate transmitter positions using the optical center of the first transmitter as the local origin. To register the coordinate measurement system to a globally referenced site frame, the locations of the transmitters must be known in the site frame.

4.2 Coordinate Frames & Transform

Two coordinate frames exist – a local frame defined by the location of the transmitters and a globally-reference site frame. To register the local frame to the site frame, the transmitters are placed on benchmarks surveyed in the Maryland State Plane Coordinate System of 1983 (MD SPCS) using the North American Datum 1983 (NAD83) as the horizontal datum and the North American Vertical Datum 1988 (NAVD88) as the vertical datum.

A transform matrix was developed to translate (move a point in space a finite distance along a given vector direction) and rotate the Cartesian points measured in the local frame to the Geodetic-Latitude-Longitude-Ellipsoid height of the site frame. Five points known in both the local and site frames are necessary to perform the transform: transmitter locations and three additional reference points. Figure 3 shows the relative location of the five points for a local frame, “Site Frame 1”, established for the implementation. K15 and K12 are used to establish a local X coordinate axis with positive direction from the origin transmitter K15 to the reference transmitter K12. The Y-axis is established perpendicular to the X-axis by using the right hand rule around the vertical Z-axis with positive direction upwards. Three other non-collinear points identified by K13, K16 and K19 serve as reference. These last three points must be non-collinear to provide a unique solution for the transform matrix. Refer to Figure 4 for a visual guide of the vector transform used for the implementation.

4.3 Part Locator Routine

A part locator routine was developed to visualize the final part position and orientation in the site frame. Developed in JAVA, the simple, robust part locator algorithm runs on the client-side and computes the part pose from the three measured fiducial point locations in the site frame. The algorithm takes the three inputs associated with a steel part from the transform routine and fixes the first point in the site frame, uses the second point to define the axis and the third point to define the plane. While this is not a precise method for determining the part pose, it does serve the purpose of enabling the visualization of the part within the site frame and for providing an estimate of the as-built location of a part's final pose.

5.0 INTEGRATION & COMMUNICATION

The key to successful development of the prototype implementation resides in seamless communication between the separate components. Component tracking data from both the ArcSecond measurement tool and the bar code scanner is captured, reformatted, and stored in the wearable computer's keyboard buffer via NIST-generated utilities on the wearable computer. These data are then displayed through a series of web pages and stateless CGI programs. The component-tracking library, linked through the CGI programs, queries the remotely hosted PIMS for related part information such as VRML models. Finally, CORBA-compliant communication routines provide the interface between the component-tracking library and the PIMS. Figure 5 illustrates the various communication methods developed for the implementation.

5.1 Wireless Communications

Running CAT-5 (Ethernet) cable across an active construction site is impractical so it is necessary to setup a wireless network. This allows users on the site to use untethered computers to communicate with other computers on or off the site. The Orinoco wireless networking hub and PC cards by Agere Systems were used to support TCP/IP networking to connect the wearable computer to the server.

6.0 CONCLUSIONS

The prototype implementation demonstrates the successful integration of various sub-systems and applications using standard interface protocols for the automated tracking of structural steel members at a job site. Enabling the transfer of real-time on-site metrology and metrology-based data from the job site to the appropriate applications and users supports operations that are dependent upon knowing and communicating the exact location and orientation of objects on the construction site.

Some areas for improvement and future work relating to the prototype system are discussed in the following two sections.

6.1 Accuracy of Reported Pose

The accuracy of the measured 3-D coordinate data supplied to the part locator routine should be quantified to provide an accuracy metric with the reported pose. This metric will enable a more precise determination of the accuracy of the reported pose for the part – a necessary component for the tracking system to be truly useful for operations such as the automated placement of structural steel members to within standard tolerances.

6.2 Communication

While the system communicated the essential information from the field to the project information management system, there are several areas to expand for error handling, error recovery, and extensions to the communications infrastructure.

As this was a prototype system, error handling and recovery were only minimally implemented. However, in real systems, error handling will need to handle users errors such as setup and registration of the coordinate measurement system, incorrect matching of field measurements to objects, and collection of collinear fiducial points. These are places where people interact with the system to provide information, and thus are potential sources of error that must be handled reliably.

The communication infrastructure, while complete for the given application, needs to be extended for future use. First, define a grammar for field measurements that can be common to all programs that use field measurements. Second, establish a communication protocol for coordinate system transforms since the exchange of the transforms between systems is not handled well in the prototype system. Third, express the system in as neutral a format as possible, so as to be extensible to future communication technologies.

7.0 REFERENCES

[1] Furlani, K.M., and Stone, W.C., "Architecture for Discrete Construction Component Tracking," Proceedings, 16th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction, Madrid, Spain, September 22-24, 1999.

[2] Stone, W.C., Furlani, K., and Pfeffer, L., "Automated Part Tracking on the Construction Site," Proceedings, 4th International Conference on Robotics for Challenging Situations and Environments (Robotics 2000), Albuquerque, NM, Feb 27 – Mar 2, 2000.

[3] Furlani, K.M., and Pfeffer, L.E., "Automated Tracking of Structural Steel Members at the Construction Site", Proceedings, 17th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction, Taipei, Taiwan, September 18-20, 2000.

[4] DeWitt Latimer IV, "Work Towards an Application Protocol to Support Interoperatbility of Sensors in Construction Metrology (LiveView)", National Institute of Technology Internal Report (NISTIR) 6867, National Institute for Standards and Technology, Gaithersburg, MD, USA, May 15, 2002.

[5] Pfeffer, L.E., and Latimer, D.T., "Toward Open Network Data-Exchange Protocols for Construction Metrology and Automation: *LiveView*," Proceedings, International Conference on Automation and Robotics in Construction (ISARC-99), Madrid, Spain, September 22-24, 1999.

[6] CIMSteel Integration Standards, Release 2, URL: http://www.cis2.org, The Steel Construction Institute, Ascot, Berkshire, UK, URL: http//www.steel-sci.org.

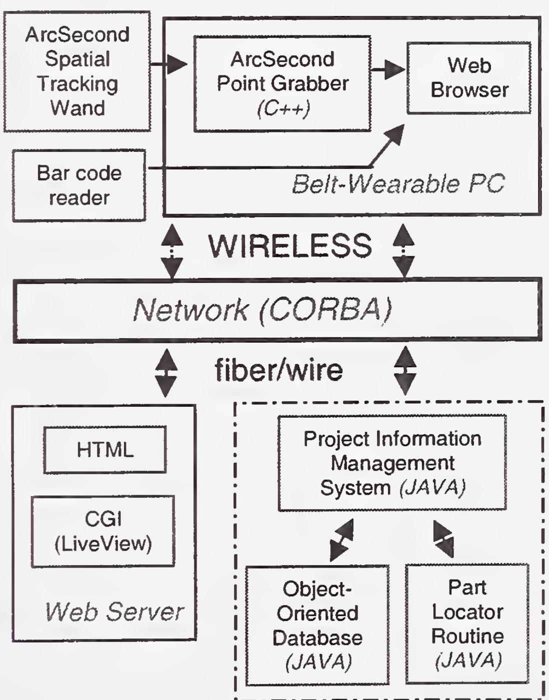


Figure 1. System architecture developed for the prototype implementation.

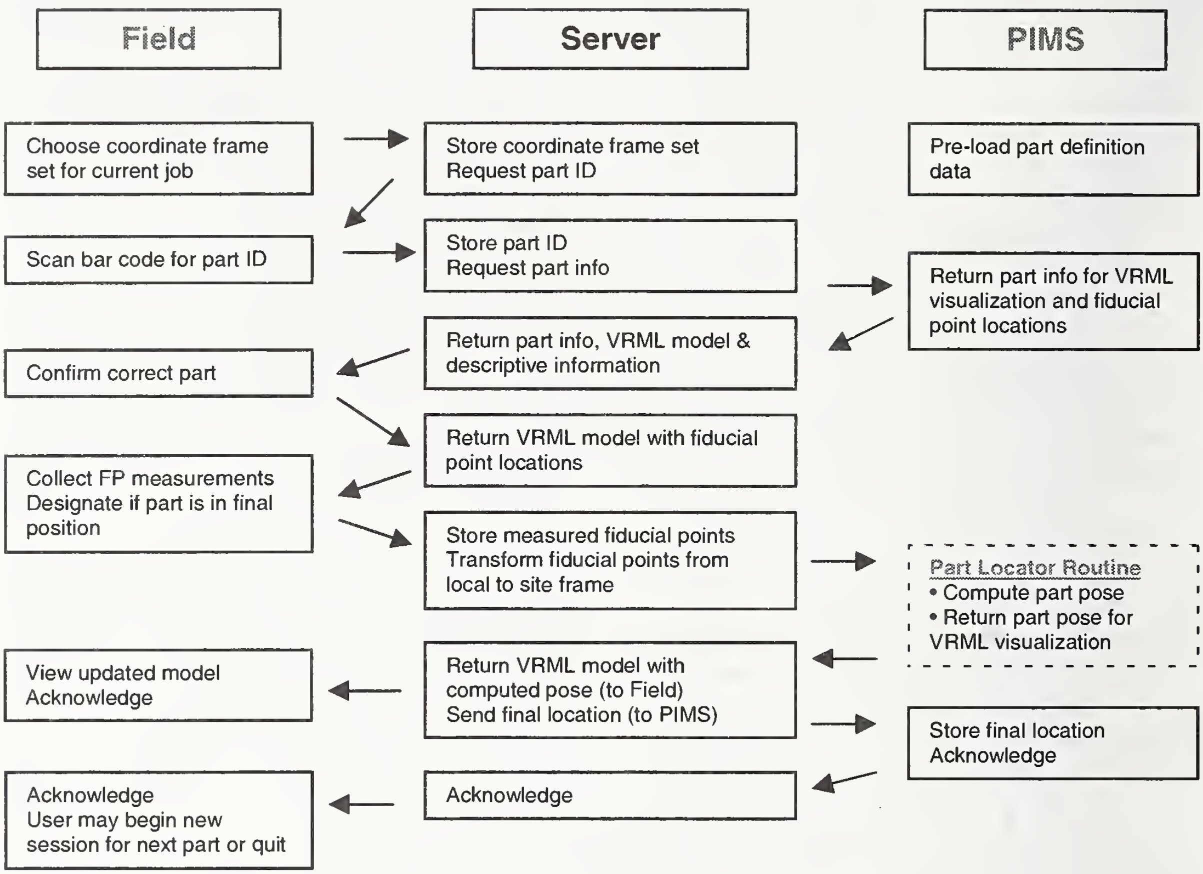


Figure 2. Operational flow chart outlining steps between the Field, Server and PIMS

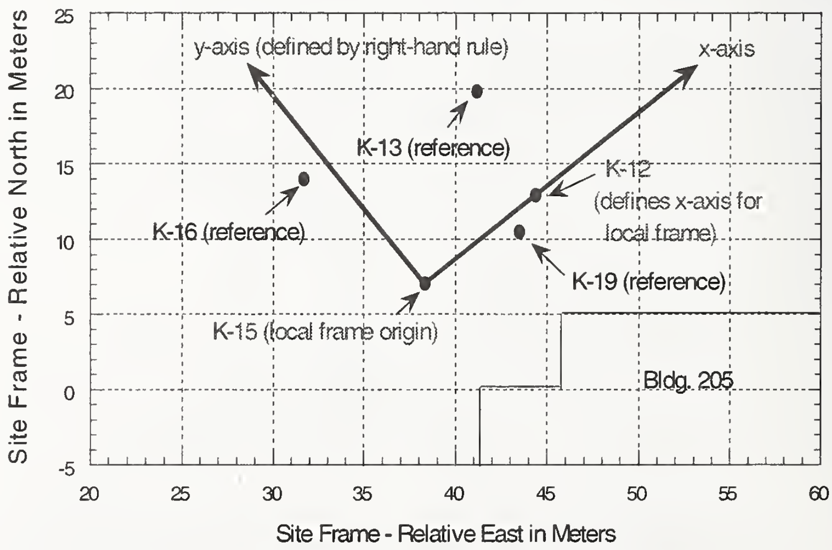


Figure 3. The local frame graphed within the site frame.

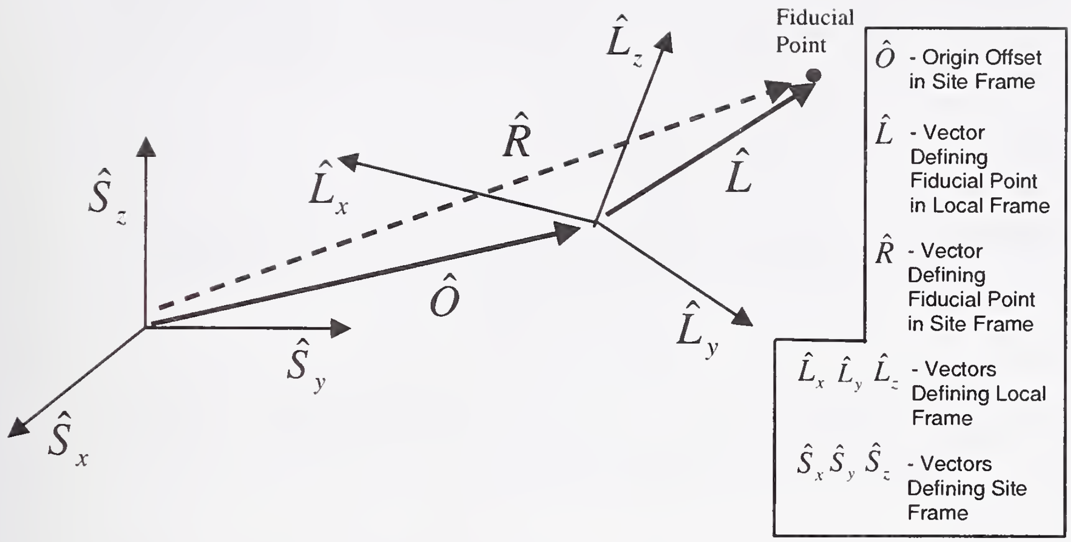


Figure 4. A vector transform from the local frame to the site frame.

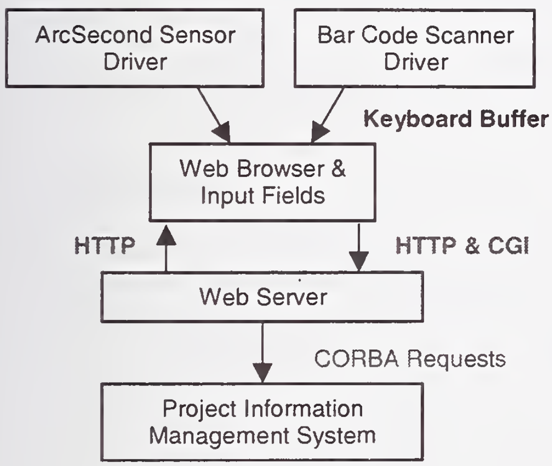


Figure 5. Communication methods used between system modules

THE EFFICIENCY OF A 3-D BLADE CONTROL SYSTEM IN THE CONSTRUCTION OF STRUCTURE LAYERS BY ROAD GRADER – AUTOMATED DESIGN-BUILD OF ROAD CONSTRUCTION IN FINLAND

Rauno Heikkilä

Dr. Tech., Senior Research Scientist, University of Oulu, Faculty of Technology, Research of Construction Technology, Finland, http://www.rate.oulu.fi, rauno.heikkila@oulu.fi

Mika Jaakkola

MSc. Tech., Research Scientist, University of Oulu, Faculty of Technology, Research of Construction Technology, Finland, http://www.rate.oulu.fi, mika.jaakkola@oulu.fi

Abstract: The paper reports on the research work in the domain of automated road construction. A new method and prototype of automated road grader has been developed in Finland. The working experiments show measurable influence and quality as well as economic profit to be achieved by the new technology.

Keywords: accuracy, automated construction, construction economics, design-build, economic efficiency, road grader, 3-D machine control

Introduction

Fast progress of teleinformatics, positioning technologies and 3-D-road design software has enabled to develop new blade control systems for road construction machinery. The major expectations of these systems are to increase productivity of the construction project and prefer quality of the end product. The objective is to build a complete digital link from design to layout. With the installation of centimeter accuracy guidance on board the earthmoving machines, the operator can fill and cut surfaces in relation to the road design. The goal of this development project is very challenging but if work is successful and reach objectives, economic benefit can be outstanding.

According to the publications of Finnish Road administration the capacity of traditionally controlled (without 3-D-machine control) motor grader on road finishing work is in practice about 7000 m^2/shift. While grader is doing for both spreading and finishing work method capacity is in normal road works about 2300 m3itd/shift (Work planning information for Road Construction 1-2, TVH).

Automatizing the total process of road construction

Since late 1990 there has come several internationally marketed motor grader 3-D machine quidance and control applications developed by such companies as Trimble, Leica and Topcon with their partners. In these applications, blade positioning needed is based on robot tacheometer or RTK-GPS. In Europe, some countries have also developed their own systems for 3-D machine control. Detailed research results about economic benefits and other experiences have very little been introduced and thus not available.

Total numerical working process of road construction consists of four part-systems. Helicopter laser scanning measuring system measures first terrain model. Measuring helicopter is positioned with GPS technology. Measuring data, terrain model which accuracy is 5-20 centimeters, is used for road structure design process. To produce steering data for the blade of construction machine we need 3-D road design software to produce 3-D design plan. In addition, the design plan has to translate to actual machine control data. (Fig. 1).

In the construction machine there must be a "comparing-software" and blade controlling system unit, which compares the real time positioning data with the 3-D road-design and calculates the steering data for the blade controlling system. In the end of the cycle the blade control system steers blade to right elevation and slope. The construction machine and the blade must fit on with the slope-sensors to pinpoint the blade position in proportion to construction machine (picture). Controlling system includes also real time quality control system of as-built road structure. As-built height and slope accuracy and efficiency on work process is documented real time. Time and money can be saved even 50 % by reducing measuring costs and by more efficient grader at road construction site.

The object of our research and development project "Intelligent road construction" is set to develop a new numerical operation process for road construction. Measurements will be operated by advanced 3-D measurement technologies and computer aided design by 3-D CAD tools. 3-D design models will be control directly without delay automated measuring and construction machinery. Finally, the quality control will be executed through automated quality control tools.

In this paper, the first result of the intelligent road construction program – the automated road grader and its working experiments are described. The construction of the mecathronics of the road grader is solved by integrated total solution principle. Hence, the movement possibilities of the different parts of the machine are very extensive.

A new integrated total solution principle and the first practical implementation has been developed in Finland. In this system the sensors and developed software algorithms control all of the directions of blade's movements. Developed system controls blade's height, slope and driving line automatically but also gives the operator possibility to control all blade movements concurrently by one joystick. This function gives to an operator a possibility to control real time the movements of gravels. Robotic total stations are used as a positioning system. When operator opens the system it automatically starts to control robotic total station.

Furthermore the machine control system contains integrated quality control system for final height and slope of the road layer and also graders capacity measuring features. The control system is working in PC and thus the operator can utilize wireless data communication, which is also included in system. Road models, quality control and capacity data are carried forward by email or in addition by data card.

The 3-D system has graphical interface from where the operator can see blade position compared to design. The digital design of the road layer is created with Terra Street design software. New company Roadsys Ltd has started to market the 3-D blade control system. The trade name is Robot Grader (Fig. 2).

Site tests

Two experiments were executed in two road construction sites as parts of motorway construction –projects in Tornio-Kemi and Salo in Finland, 2000-2001.

Results

The detailed results of the work studies executed are introduced in tables 1-5. In the tables, the term "basic capacity" is determined to be machines work achievement per basic working time, which is the time when machine is all time working. The term "method capacity" means machine's work achievement per method working time, which includes basic time and other time used in efficient working for example setting times of tacheometer, waitings of truckers, etc.

According to the tests (table 1) the standard deviation of the 3-D blade control system in unbound base course is about 17 mm in height direction. According to the specifications of Finnish Road Administration the tolerance for that work phase and product part is ± 20 mm. All of the slope measurements (table 2) fulfill the tolerance requirements. According to the work studies (tables 3-4) the economic benefits grew best up more than 50 %. It is also essential to notice that the needs for measurement crew as well as the costs are decreased.

Conclusion

The practical functionality of the operation principle was verified by the experiment results.

The adequacy of accuracy was proven to be sufficient. Economic benefits by automation were observed.

The road and traffic design and construction techniques are developing rapidly worldwide. Teleinformatics and positioning technologies play important part in this process. Finland is one of the countries in developing these new technologies and also suitable operating environment for testing them. Construction industry should without prejudice start to use and take advantage of these new technologies.

Intelligent Road Construction Site IRCS is a Finnish research and development programme been started in Finland. The programme consists of several different research and development projects. The entirety is scheduled 2001-2005 and it collects about twenty infra companies and other parties of road construction process. The total investment for research activities is about 2 M€.

ACKNOWLEDGEMENTS

The authors want to warmly thank VTT Elektronics, Roadsys Oy, Terrasolid Oy, Geotrim, Patria Vammas, Tieliikelaitos and TEKES – the National Technology Agency in Finland. The project belongs as a part to the Intelligent Road Site –research & development – programme executed in Finland.

REFERENCES

Heikkilä, R. (1996) The Applicability of an Electro-Optical Coordinate Measuring Technique to the Dimensional Control of Precast Concrete Facade Panel Production. Oulu, Acta Universitatis Ouluensis, Technica C 86. Doctor's thesis. 87 p. + Appendixes.

Heikkilä, R. & Jaakkola, M. (1999) New Automation Possibilities for Road Construction Methods. Košice, the 3rd Scientific Conference with International Participation "Building and Energy" in Košice, 8-9 September 1999.

Heikkilä, R. & Jaakkola, M. (2000) Planning Future Activities for Road Construction Automation Purposes. Construction Economics News, 1/2000, p. 6-9.

Jaakkola, M. & Heikkilä, R. (2000) Liikkuvan työkoneen ohjaus tienrakentamisessa. Tiehöylän ja murskeenlevittimen CAD/CAM-järjestelmät. Tienrakennusautomaatio. (Controlling mobile work machines in Road Construction. CAD/CAM systems for road grader and mass spreader. Automated Road Construction. Project Report. Confidential. Oulu, University of Oulu.

Jaakkola, M. & Heikkilä, R. (2001) Automatizing a road grader for Finnish Road Company – Development Phase I. Project Report. Confidential. Oulu, University of Oulu.

TVH 733892 (1983). Tien tekemisen työnsuunnittelutiedot 1-2 (Work planning information for Road Construction 1-2). Finnish Road Administration.

SOURCES FOR PRODUCT INFORMATION

www.leica-geosystems.com
www.topconlaser.com/MC_Products
www.trimble.com/bladepro3-D.html
www.mikrofyn.com
www.precisesurvey.com
www.roadsys.fi

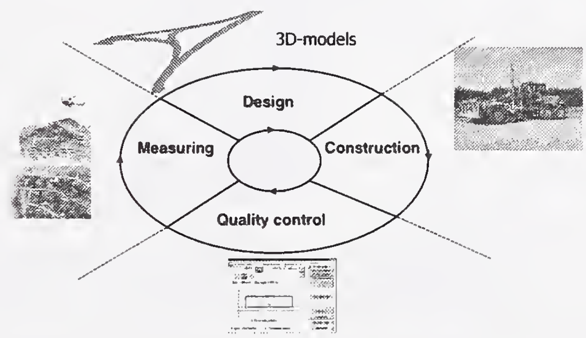


Figure. 1: The principle of the numerical working process of road construction.

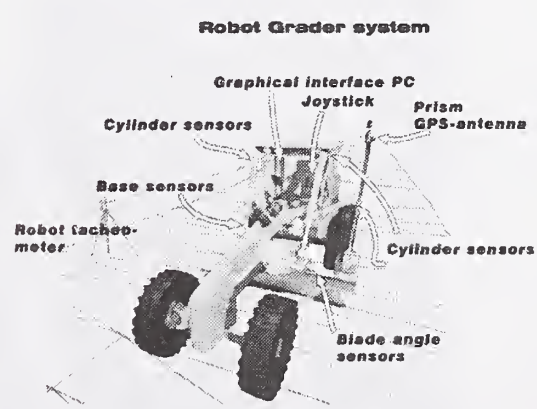


Figure. 2: Robot grader. A new 3-D blade control system principle for road grader.

Figure. 3: The 3-D surface model created for the tests. The model was designed by Terra Street CAD application and used directly to control the blade of grader.

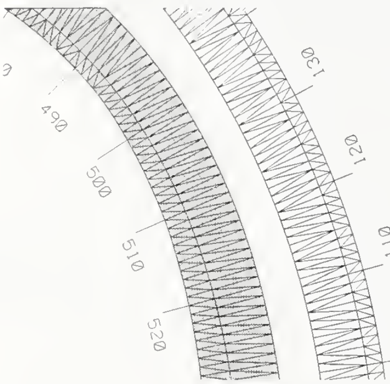


Figure 4: An example of 3-D surface model of Tornio-Kemi motorway.

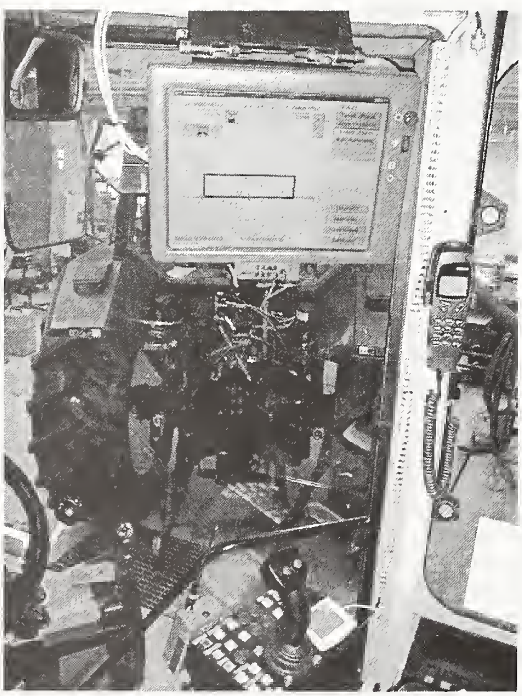


Figure 5: The user interface for the 3-D blade control system with wireless data communication system.

Figure 6: The site tests of the 3-D blade control system.

Table 1. Automated road grader experiments. The height results of Tornio-Kemi site tests.

Site number	Average [mm]	St.dev. [mm]	Min [mm]	Max [mm]	Count [number]	Accepted [%] *)
E5R1	- 23	19	-64	22	33	66
E5R3 Jakava	-31	19	-60	19	27	50
E5r4	-19	17	-53	22	48	81
E5R3 Kantava	-21	16	-59	10	36	61
MO plv 16860-17780	-8	17	-50	41	146	87
MO plv 15620-16240	-3	10	-30	10	90	99

*) Road structures vertical position tolerances were ± 20 mm

Table 2. Automated road grader experiments. The slope results of Tornio-Kemi site tests.

Site number	Average [%]	St.dev. [%]	Min [%]	Max [%]	Count number	Accepted %
E5R1	-	0,26	-0,4	0,4	15	100 %
E5r4	-	0,56	-0,8	1,2	14	85 %
MO 16900-17780	-3,07	0,25	-0,6	0,5	46	100 %

Table 3: Automated road grader experiments. The capacity results in adjustment tasks.

Site number	Basic [m2/h]	Method [m2/h]	Area of object [m2]	Adjusted aggregate [t]
mo plv 18760-19020	1282	803	3120	-
mo plv 18600-18820	2258	1650	2860	-
E5r4	2444	2005	2240	18
E5R1	715	677	2100	71

Table 4: Automated road grader experiments. The capacity results in grade tasks.

Site number	Basic [m2/h]	Method [m2/h]	Method of work [t/h] / [itdm3/h]	Adjusted aggregate [t]
E5R4 jakava	617	579	162 / 108	590
E5R3 kantava	368	308	238 / 159	1358
mo plv 17360-17580	550	491	217 / 145	989
mo plv 17200-17460	1613	1366	747 / 498	1445
mo plv 17920-18600	1678	1498	87 / 58	457

OPTIMAL CONTROL OF AN EXCAVATOR BUCKET POSITIONING

by

E. Budny, M. Chłosta, W. Gutkowski [1)]

ABSTRACT.
Recently, there is an increasing interest in controlled excavation processes. However, the main attention, in research works, is paid to the bucket motion. This part of the process can be considered as a quasi static, kinematically induces process [3]. It means that dynamic effects, by dropping accelerations terms can be neglected. This is not a case in the second part of the process consisting of: lifting the bucket filled with the soil, swinging the whole excavator with respect to vertical axis, lowering the bucket and discharging it. Next, the bucket is brought back to the excavation place again. Discussing these motions, one has to taking in to account dynamic effects. It should be also noted that mentioned motions are lasting approximately the same time as the digging process. It is then worthy to try to minimize the time needed for bringing the filled bucket to the discharge place, and back to the digging site. It is then the aim of the paper to present an optimal control of such a minimum time process. The paper deals with an optimum problem of positioning an excavator bucket along prescribed trajectory using minimum time. The paper is illustrated with numerical results giving some optimal trajectories.

KEYWORDS: excavation, optimization, control.

1. INTRODUCTION

A standard excavation process, on the construction side, can be seen as composed of two parts. The first one is the process of digging and filling the bucket with soil or other material. The second part of the operation consists in lifting (L) the filled bucket, swinging (S) it with respect to a vertical axis, stopping (S) it at the place where it should be unloaded, and discharged (D). The whole process is below defined as LSSD.

Since about twenty years, many attention has been paid to robotics application in the construction industry. However, most attention has been paid to the digging processes. Relatively large number of works in this fields were presented at the International Symposia on Robotics in Construction. Among others, Budny *et al.*[3] proposed an load-independent control of excavation process along a prescribed trajectory. The idea of the paper was to propose a control system free of a number of sensors mounted on the excavator attachment.

The present study is dealing with the LSSD process, namely with the problem of minimum optimal positioning of excavator bucket from the position where it is filled with the soil, ending the digging, to the place where it should be discharged. There are several important reasons to undertake this research. They can be listed as follows:

(1) LSSD is taking often as much time as digging process;
(2) automation of LSSD would decrease the operator efforts, comparing with hand controlled motion of the bucket;
(3) optimization of LSSD should decrease the time and/or energy needed for its realization.
(4) automation of LSSD should decrease the probability of accidents, due to human errors.

It should be noted that available modern software and hardware, with their decreasing cost, make practical realizations of discussed positioning possible. The 3D visualization on the screen of a monitor would allow an operator to locate the bucket position. Then with a “push button” command, supported by an appropriate algorithm, would allow to automatic motion of the bucket along prescribed trajectory to discharging place.

The positioning and optimal control, have found also interest in pneumatic and hydraulic systems applied in machines for construction industry. Rachkov et al. are considering optimal control of a pneumatic manipulator with performance index imposed on energy con-

[1)] Institue of Mechanised Construction and Rock Mining, Poland, 02-673 Warsaw, ul.Racjonalizacji 6/8, e-mail: mch@imbigs.org.pl

sumption. Shih et al. propose to apply sliding mode method, in positioning of a pneumatic cylinder with high speed solenoid valves.
Starting this work, it was essential to assume an appropriate algorithm based on a rigorous mathematical back ground for proposed trajectory optimization (TO). The first theoretical results in this field are due Pontryagin [5]. In his famous mathematical theory of optimal control the basic concepts of TO are presented. An application of the Pontryagin principle to a two arm manipulator is the work by Avetisian et al. [1].
Almost the same time with Pontryagin theory, emerged the non-linear programming (NLP) with its basic theorems by Kuhn-Tucker. Its development has been strongly dependent of digital computers, and then gave very powerful tool to solve many problems with discretized variables. In a survey paper by Betts [2] application of NLP and of some other methods related to TO are presented. Recently, Furukawa [4]proposed a trajectory planning dicretizing functions entering in the problem, into piecewise constant functions. Roh and Kim [6] are proposing an indirect, applying time FEM method, combined with NLP. In the proposed algorithm is taking some elements of the both, last mentioned papers. The curve, joining initial and final bucket positions, is divided into a given number of equal elements. Unknowns in NLP problem are times needed to travel along a segment with a constant velocity. The system of equations and inequalities arising from NLP problem are solved by successive approximations method. In the first step, some independent variables are assumed, and other solved from state equations. Next, the remaining equations are solved from state conditions equations, giving values for next approximation. Some illustrative examples are presented at the end of the study.

2. KINEMATICS AND DYNAMICS

Consider a simplified model of an excavator with a loaded bucket of mass M_1. Simplification consists in assumption that all three members of the excavator attachment constitute one solid beam of length 2L and mass M_2. The arm, driven by a hydraulic actuator, can rotate with respect to a horizontal axis, by an angle α (Fig. 3). Additionally the arm, with the bucket, can rotate with respect to a vertical axis by an angle φ. The latter motion is due a hydraulic motor rotating the carriage.
The considered system is then of two degrees of freedom. Its kinetic energy, a function of two unknowns angular velocities $\dot{\alpha}$ and $\dot{\varphi}$ is equal to:

$$T = \frac{1}{2} I \left(\dot{\alpha}^2 + \dot{\varphi}^2 \cos\alpha \right) \tag{1}$$

where:

$$I = 4\left(M_1 + \frac{4}{3} M_2 \right) L^2$$

The kinetic energy reaches its value from work V exerted by external forces:

$$V = MgL \sin\alpha \tag{2}$$

where:

$$M = 2M_1 + M_2$$

From the second order Lagrange equations we get the following equations of motion:

$$\begin{aligned} I\left(\ddot{\alpha} + \frac{1}{2}\dot{\varphi}^2 \sin 2\alpha \right) &= Q_1 - MgL\cos\alpha \\ I\left(\ddot{\varphi}\cos^2\alpha - \dot{\varphi}\dot{\alpha}\sin 2\alpha \right) &= Q_2 \end{aligned} \tag{3}$$

The system is then transferred in to a set of four equations of first order each, by assuming the following variables

$$\left[\alpha, \varphi, \dot{\alpha}, \dot{\varphi} \right] = \left[x_1, x_2, x_3, x_4 \right] \tag{4}$$

$$\begin{aligned} \dot{x}_1 &= x_3 \\ \dot{x}_2 &= x_4 \\ \dot{x}_3 &= \frac{1}{I}\left[-\frac{1}{2} I x_4^2 - MgL\cos x_1 + Q_1 \right] \\ \dot{x}_4 &= 2x_3 x_4 tg x_1 + \frac{Q_1}{I\cos^2 x_1} \end{aligned} \tag{5}$$

3. PRE-SHAPED INPUTS FOR MINIMUM TIME CONTROL

Our strategy now is related to minimum time needed for the bucket to travel from an initial position α_0, φ_0 to a final position α_f, φ_f. We are then discussing a pre-shaped function for control input, however this is related to an open-loop control. Only after finding this function it would be possible to design a close loop system.

The discussed problem is stated as follows:

Find torques Q_1 and Q_2, driving the arm with respect to the horizontal axis, and with respect to vertical one, assuring the shortest time for the bucket to move from the initial to the final position along a given curve. The problem can be stated in terms of formulate in the form:

Find the minimum time

$$t_f - t_0 \to min \tag{6}$$

to move the bucket from initial point α_0, φ_0 and initial velocity $\dot{\alpha}_0 = \dot{\varphi}_0 = 0$, to the final position α_f, φ_f, and final velocity $\dot{\alpha}_f = \dot{\varphi}_f = 0$

if the motion of the bucket is defined by state equations (5) and torques are Q_1 and Q_2 are bounded as follows:

$$\begin{aligned} -Q_{1,0} \le Q_1 \le Q_{1,0} \\ -Q_{2,0} \le Q_2 \le Q_{2,0} \end{aligned} \tag{7}$$

The equations (5), (6) and (7) represent a nonlinear optimization problem, which can be solved numerically only. We start with discretization of variables entering into the problem. The given trajectory between the starting bucket position α_0, φ_0 and its final position α_f, φ_f is divided into j_o elements. The traveling time $t_f - t_0$ is then divided in j_o time intervals of $h = \frac{t_f - t_0}{j_0}$ each.

The state equations (5) after discretization take the following form:

$$\begin{aligned} & I\frac{x_{i,j+2} - 2x_{i,j+1} - x_{i,j}}{h^2} + \\ & -\frac{1}{2}I\,sin\left(2x_{i,j}\right)\frac{\left(x_{i,j+1} - x_{j,i}\right)^2}{h^2} + \\ & +MgL\,cos\,x_{i,j} = Q_1 \\ & I\,cos\left(x_{i,j}\right)\frac{x_{i,j+2} - 2x_{i,j+1} + x_{i,j}}{h^2} + \\ & -I\,sin\left(2x_{i,j}\right)\frac{\left(x_{i,j+1} - x_{i,j}\right)\left(x_{i,j+1} - x_{i,j}\right)}{h^2} = Q_2 \end{aligned} \tag{8}$$

The following algorithm is proposed to solve the discussed problem.

Step1. Assume lower bound for h: $h_l = 0$. In this case constraints (7) are not fulfilled as torques would be infinitely large.

Step 2. Assume upper bound for h: h_u, assuring that torques found from state equations (8) fulfil constraints (7).

Step 3. Take $h = \frac{h_l + h_u}{2}$.

Step 4. Solve torques Q_1 and Q_2 from state equations (8) and verify (7).

If constraints (7) are fulfilled, substitute for $h_u = h$.

If constraints (7) are nor satisfied, substitute $h_l = h$.

Step 5. If $h_u - h_l \ge e$, where e is assumed admissible error, go to Step 3.

If $h_u - h_l < e$ then STOP.

4. EXAMPLES

Assume the following data for values entering into the problem:

$M_1 = 440\,[kg] \quad M_2 = 160\,[kg]$

$M = 1040\,[kg]$

$I = 6500\left[kg \cdot m^2\right]$

$L = 1.500\,[m]$

$-2000 \le Q_1 \le 2000\,[N \cdot m]$

$-6000 \le Q_2 \le 6000\,[N \cdot m]$

For the data specified above, two distinct situations are considered.

Case 1:The initial bucket position is $x_{1,0}$=-30^o, $x_{2,0}$=0^o

Its final position is $x_{1,f}$=60^o, $x_{2,0}$=120^o

At both positions, the bucket velocity is equal to zero. The assumed traveling curve, represented by angles α and φ are given in Fig. 4 and Fig. 5. The final result, showing relations between torques Q_1, Q_2 and time are given in Fig. 6, Fig. 7.

Case 2: The initial bucket position is the same as in Case 1.

The assumed traveling curve, represented by angles α and φ are given in Fig. 8 and Fig. 9.

The final results, showing relation between torques Q_1 and Q_2 and time are given in Fig. 10 and Fig. 11.

5. CONCLUSIONS

An optimization problem for minimum traveling time for an excavator bucket, between along given trajectory is presented. Numerical results show significant differences, which may take place, between torques/time relations for two different trajectories.

The problem extended to a real excavator with three degrees of freedom, could be easily implemented into control of LSSD processes in serial manufactured excavators.

6. ACKNOWLEDGEMENTS

The support of Polish State Committee for Scientific Research, grant nr 8 T07A 035 20 is gratefully acknowledged.

7. REFERENCES

[1] Avetisian, V.V., Bolotnik, N.N., Cherneusko, F.L." Optimal Programmed Motion of a Two Arm Manipulator" (in Russian), Tech. Kibernetika, No.3, 1985

[2] Betts,J.T.,"Survey of numerical methods for trajectory optimization", J. of Guidance ,Control, and Dynamics,Vol.21,No2, March-April, 1998.

[3] Budny E.,Chłosta M.,Gutkowski W. „Experiment on a stable motion of the backhoe excavator bucket" Proceedings of the18th ISARC Conference, Krakow, Poland.,2001

[4] Furukawa, T., "Time-Subminimal Trajectory Planning for Discrete Non-Linear Systems", Eng. Opt. Vol. 34, No3, 2002,

[5] Pontryagin, L.S., Boltyanskii, V., Gamkrelidze, R. and Mishchenko, E. "The Mathematical Theory of Optimal Processes", Intersciences, 1962.

[6] Roh, W., Kim, Y. "Trajectory Optimization for a Multi-Stage Launch Vehicle Using Time Finite Element and Direct Collocation Method". Eng. Opt. Vol. 34, No 1, 2002.

[7] Shih, M.C., Ma, M. A., „Position Control of a Pneumatic Rodless Cylinder Using Sliding Mode M-D-PWM Control the High Speed Solenoid Valves", JSME International Journal, Series C. Vol. 41. No 2.,1998

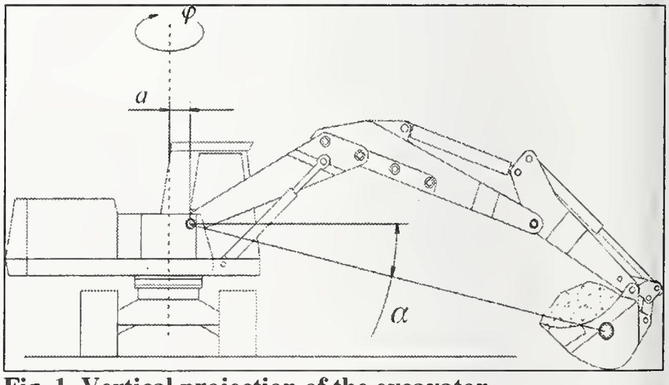


Fig. 1. Vertical projection of the excavator.

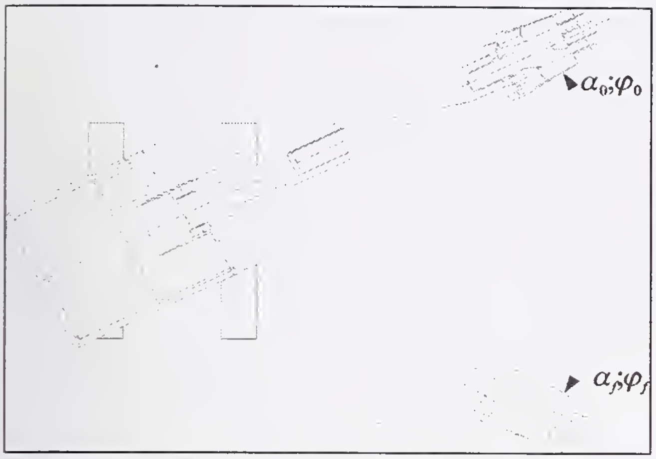


Fig. 2. Horizontal projection of the excavator

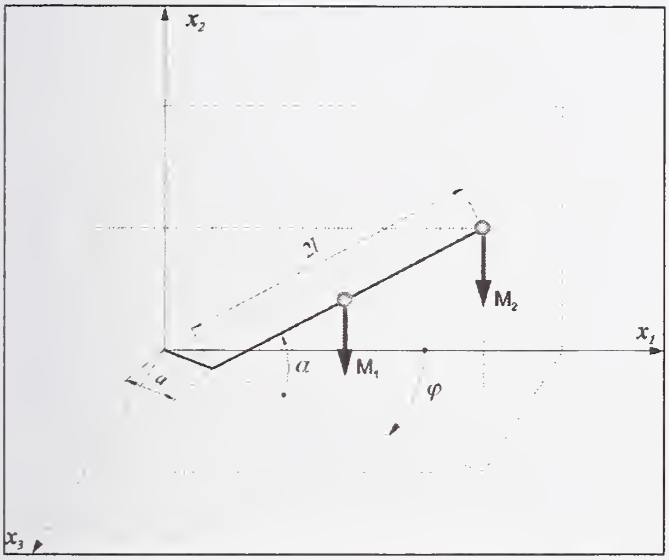


Fig. 3. Dynamic model of the excavator in coordinate system

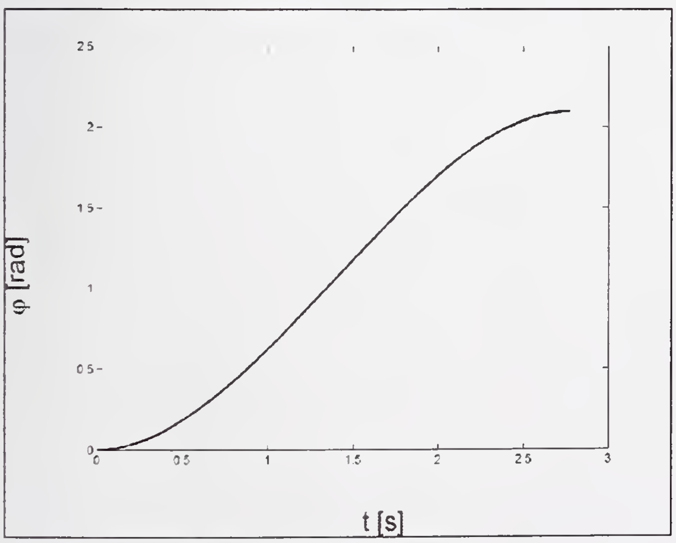


Fig. 4. Case 1: φ vs traveling time t

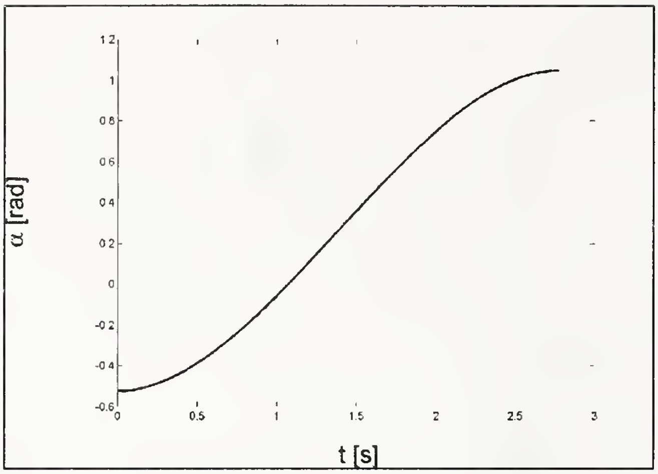


Fig. 5. Case 1: α vs traveling time t

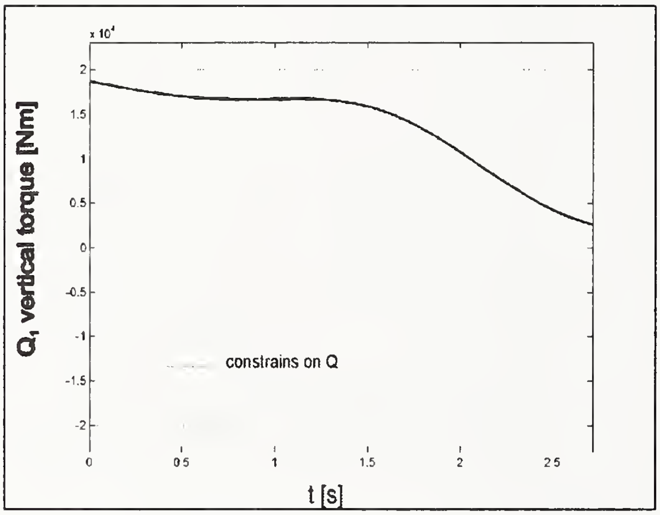


Fig. 6. Case 1: Vertical torque Q_1 vs traveling time t

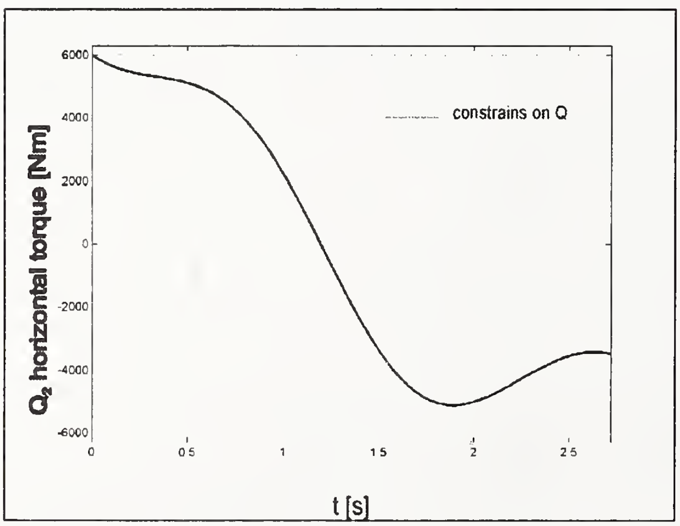


Fig. 7. Case 1: Horizontal torque Q_2 vs traveling time t

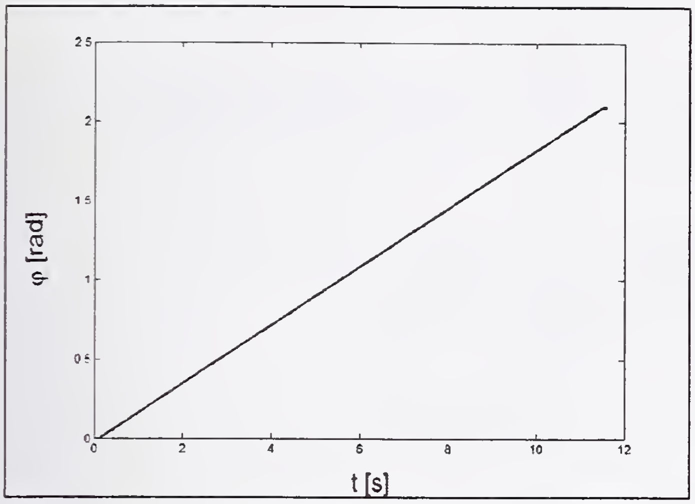


Fig. 8. Case 2: φ vs traveling time t

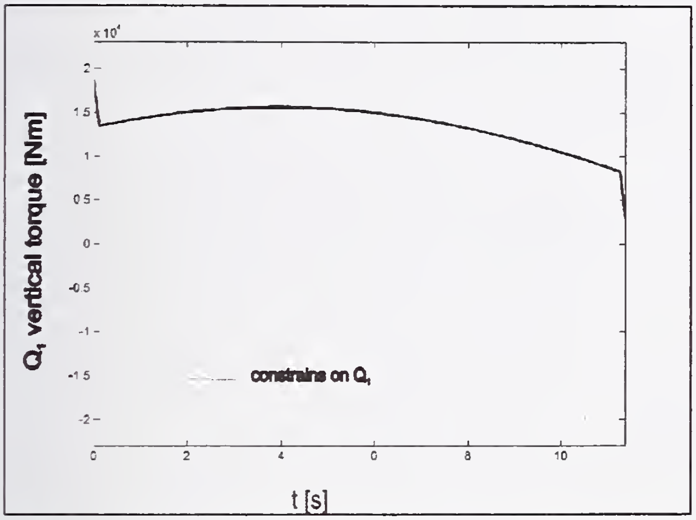


Fig. 9. Case 2: horizontal torque Q_1 vs traveling time t

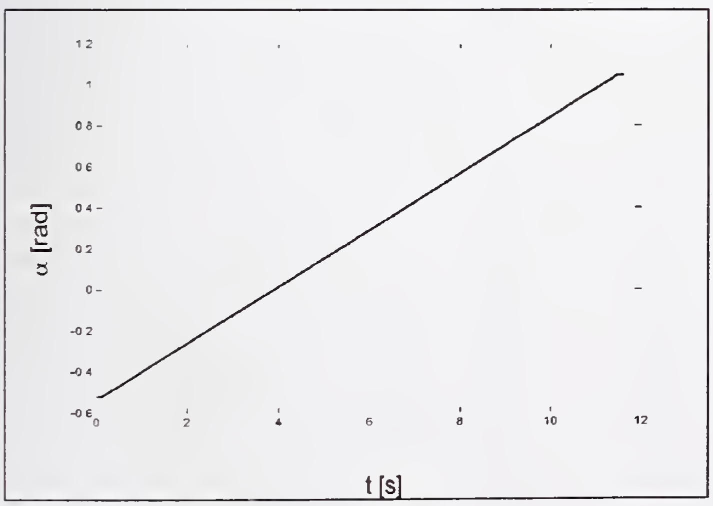


Fig. 10. Case 2: α vs traveling time t

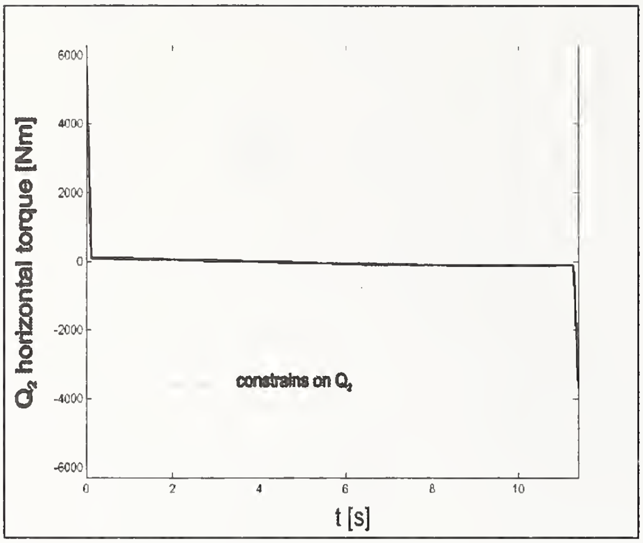


Fig. 11. Case 2: horizontal torque Q_2 vs traveling time t

Automated Construction using Contour Crafting – Applications on Earth and Beyond

by

B. Khoshnevis, Professor, Industrial & Systems Engineering
George Bekey, Professor, Computer Science
University of Southern California
khoshnev@usc.edu, bekey@usc.edu

ABSTRACT: Although automation has advanced in manufacturing, the growth of automation in construction has been slow. Conventional methods of manufacturing automation do not lend themselves to construction of large structures with internal features. This may explain the slow rate of growth in construction automation. Contour Crafting (CC) is a recent layered fabrication technology that has a great potential in automated construction of whole structures as well as sub-components. Using this process, a single house or a colony of houses, each with possibly a different design, may be automatically constructed in a single run, imbedded in each house all the conduits for electrical, plumbing and air-conditioning. Our research also addresses the application of CC in building habitats on other planets. CC will most probably be one of the very few feasible approaches for building structures on other planets, such as Moon and Mars, which are being targeted for human colonization before the end of the new century.

KEYWORDS: *Contour Crafting, housing construction, construction on other planets*

INTRODUCTION

Since the early years of the twentieth century automation has grown and prevailed in almost all production domains other than construction of civil structures. Implementation of automation in the construction domain has been slow due to: a) unsuitability of the available automated fabrication technologies for large scale products, b) conventional design approaches that are not suitable for automation, c) significantly smaller ratio of production quantity/type of final products as compared with other industries, d) limitations in the materials that could be employed by an automated system, e) economic unattractiveness of expensive automated equipment, and f) managerial issues. On the other hand, the following are reported to be serious problems that the construction industry is facing today (Warszawski and Navon, 1998):

- Labor efficiency is alarmingly low,
- Accident rate at construction sites is high,
- Work quality is low, and
- Control of the construction site is insufficient and difficult, and skilled workforce is vanishing.

Automation of various parts and products has evolved considerably in the last two centuries but construction remains largely as a manual practice. This is because the various conventional methods of manufacturing automation do not lend themselves to construction of large structures. A promising new automation approach is layered fabrication, generally known as rapid prototyping. Although several methods of rapid prototyping have been developed in the last two decades (Pegna, 1997), and successful applications of these methods have been reported in a large variety of domains (including industrial tooling, medical, toy making, etc.), currently Contour Crafting (CC) seems to be the only layer fabrication technology that is uniquely applicable to construction of large structures such as houses (Khoshnevis, 2000).

CONTOUR CRAFTING

Contour Crafting (CC) is an additive fabrication technology that uses computer control to exploit the superior surface-forming capability of troweling to create smooth and accurate planar and free-form surfaces (Khoshnevis 1998, Khoshnevis *et al.,* 2001-a; Khoshnevis *et al,* 2001-b). Some of the important advantages of CC compared with other layered fabrication processes are better surface quality, higher fabrication speed, and a wider choice of materials.

The key feature of CC is the use of two trowels, which in effect act as two solid planar surfaces, to create surfaces on the object being fabricated that are exceptionally smooth and accurate. Artists and craftsmen have effectively used simple tools such as trowels, blades, sculpturing knives, and putty knives, shown in Figure 1, with one or two planar surfaces for forming materials in paste form since ancient times. Their versatility and effectiveness for fabricating complex free-form as well as planar surfaces is evidenced by ancient ceramic containers and sculptures with intricate or complex surface geometries as well as detailed plaster work that have shapes as complicated as flowers, on the walls of rooms. Surface shaping knives are used today for industrial model making (e.g., for building clay models of car bodies). However, despite the progress in process mechanization with computer numerical control and robotics, the method of using these simple but powerful tools is still manual, and their use is limited to model building and plaster work in construction.

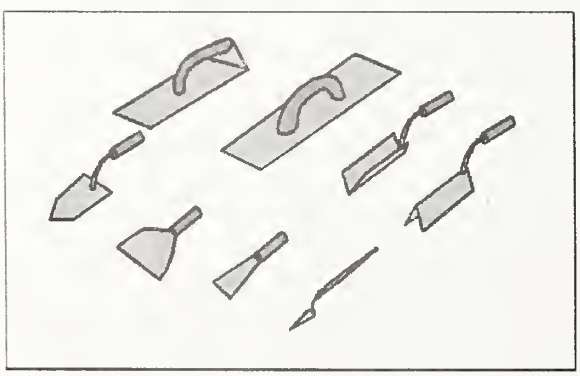


Figure 1. Simple Historical Construction Tools

In CC, computer control is used to take advantage of the superior surface forming capability of troweling to create smooth and accurate, planar and free-form surfaces. The layering approach enables the creation of various surface shapes using fewer different troweling tools than in traditional plaster handwork and sculpting. It is a hybrid method that combines an extrusion process for forming the object surfaces and a filling process (pouring or injection) to build the object core. As shown in Figure 2, the extrusion nozzle has a top and a side trowel. As the material is extruded, the traversal of the trowels creates smooth outer and top surfaces on the layer. The side trowel can be deflected to create non-orthogonal surfaces. The extrusion process builds only the outside edges (rims) of each layer of the object. After complete extrusion of each closed section of a given layer, if needed filler material such as concrete can be poured to fill the area defined by the extruded rims.

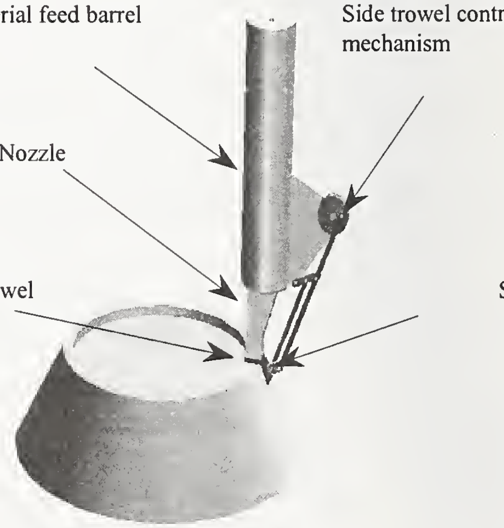


Figure 2. Contour Crafting Process

APPLICATION IN CONSTRUCTION

Aplication of CC in building construction is depicted in Figures 3 where a gantry system carrying the nozzle moves on two parallel lanes installed at the construction site. A single house or a colony of houses, each with possibly a different design, may be automatically constructed in a single run. Conventional structures can be built by integrating the CC machine with a support beam picking and positioning arm, and adobe structures, such the ones designed by CalEarth (www.calearth.org) and depicted in the Figure 4, may be built without external support elements using shape features such as domes and vaults. Following are some interesting aspects of this automated construction concept:

Design Flexibility: The process allows architects to design structures with functional and exotic architectural geometries that are difficult to realize using the current manual construction practice.

Multiple Materials: Various materials for outside surfaces and as fillers between surfaces may be used in CC. Also, multiple materials that chemically react with one another may be fed through the CC nozzle system and mixed in the nozzle barrel immediately before deposition. The quantity of each material may be controlled by computer and correlated to various regions of the geometry of the structure being built. This will make possible the construction of structures that contain varying amounts of different compounds in different regions.

Utility Conduits: As shown in Figure 5 utility conduits may be built into the walls of a building structure precisely as dictated by the CAD data. Sample sections made with CC and filled with concrete as shown in Figure 8 demonstrate this possibility.

Paint-Ready Surfaces: The quality of surface finish in CC is controlled by the trowel surface and is independent of the size of the nozzle orifice. Consequently, various additives such as sand, gravel, reinforcement fiber, and other

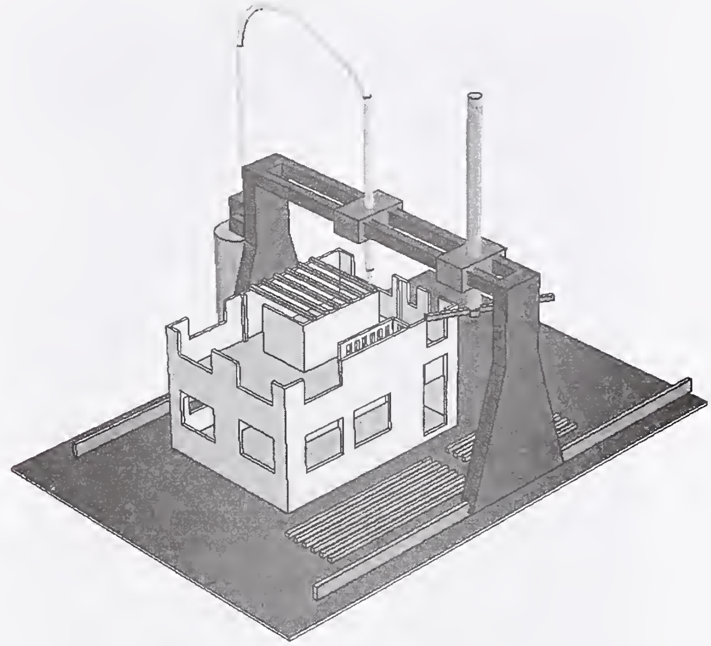


Figure 3.Construction of conventional buildings using CC

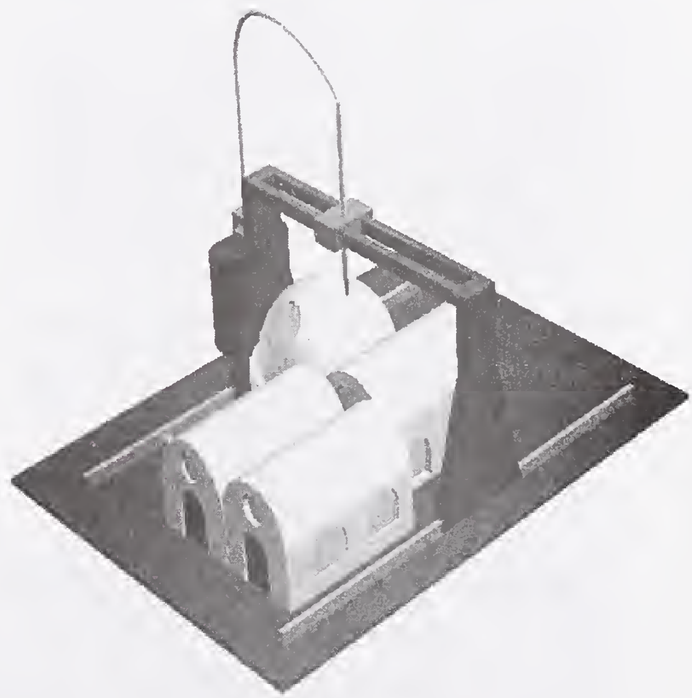


Figure 4. Construction of adobe buildings using CC

applicable materials available locally may be mixed and extruded through the CC nozzle. Regardless of the choice of materials, the surface quality in CC is such that no further surface preparation would be needed for painting surfaces. Indeed an automated painting system may be integrated with CC.

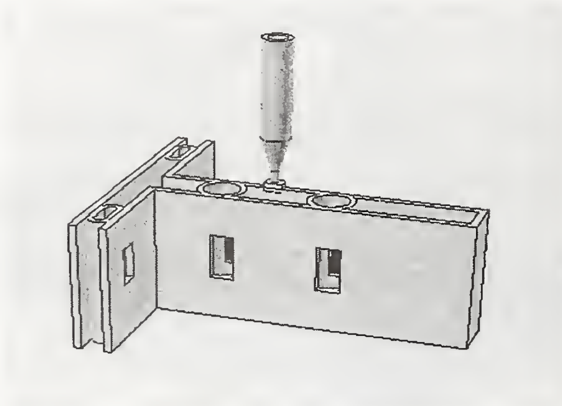


Figure 5. Complex Wall Section

Smart Materials: Since deposition in CC is controlled by computer, accurate amounts of selected construction materials, such as smart concrete, may be deposited precisely in the intended locations. This way the electric resistance, for example, of a carbon filled concrete may be accurately set as dictated by the design. Elements such as strain sensors, floor and wall heaters can be built into the structure in an integrated and fully automated manner.

Reinforcement: Modular imbedding of steel mesh reinforcement into each layer may be devised, as shown in Figure 6. The two simple modular components shown in this figure may be delivered by an automated feeding system that deposits and assembles them between the two rims of each layer built by CC. Concrete may then be poured between the rims of each layer to contain the steel mesh. The mesh can follow the geometry of the structure. Note that in this configuration the CC nozzle, the steel feeder, and the concrete filler feeder are all on the same gantry system. Such a system can create shapes with smooth outer surfaces and reinforced internal structure automatically and in one setup.

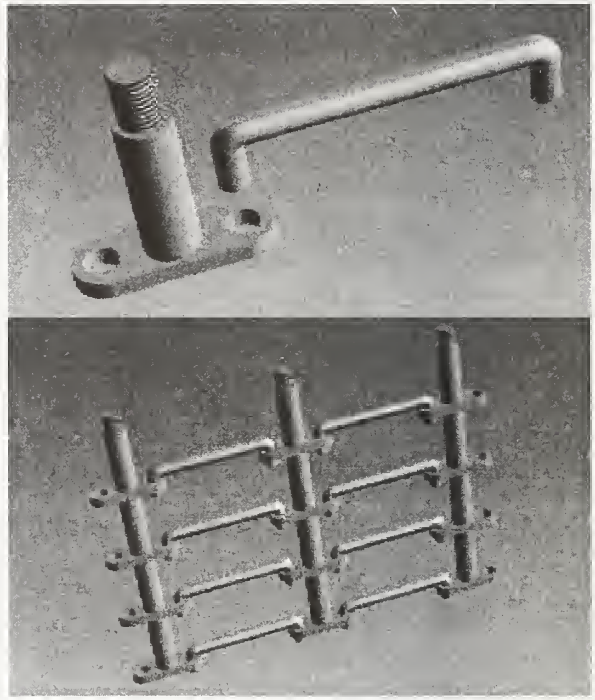


Figure 6. Steel reinforcement modules and progressively constructed reinforcement mesh

As an alternative to traditional metal reinforcement, other advanced materials can be used, such as the fiber reinforced plastics (FRP). Since the nozzle orifice in CC does not need to be very small, it is possible to feed glass or carbon fiber tows through the CC nozzle to form continuous reinforcement consolidated with the matrix materials to be deposited. In the proposed study, deposition of the FRP reinforcement by a parallel nozzle built into the CC nozzle assembly will also be considered. Co-extrusion is further discussed in a later section.

Reinforcement can also be provided using the post-tensioning system. Accurate ducts can be generated by the CC process. Similar to post-tensioned concrete construction, metal or FRP wires can be fed through the ducts and then post-tensioned to provide reinforcement.

STATE OF DEVELOPMENT AND FUTURE PLAN

Several CC machines have been developed at USC for research on fabrication with various materials including thermoplastics, thermosets, and various types of ceramics. These machine include a XYZ gantry system, a nozzle assembly with three motion control components (extrusion, rotation, and trowel deflection) and a six axis coordinated motion control system. The machine developed for ceramics processing is capable of extruding a wide variety of materials including clay and concrete.

We have conducted extensive experiments to optimize the CC process to produce a variety of 2.5D and 3D parts with square, convex, and concave features, some filled with concrete, as shown in Figures 7 and 8. The scale has been of the samples made to date (the hand in Figure 7 is indicative of the scale. We are currently working on the development of new nozzle assemblies that are especially designed for construction application. With the new nozzles we intend to first fabricate full scale sections of various building features such as sections of walls with conduits built in, and supportless roofs and perform various structural analysis and testing using a wide variety of candidate materials.

We plan to consider the NIST RoboCrane system as an alternative to a conventional XYZ gantry which may encounter problems due to rail alignment and structural rigidity.

We plan to explore the applicability of the CC technology for building habitats on the Moon and Mars. In the recent years there has been growing interest in the idea of using these planets as platforms for solar power generation, science, industrialization, exploration of our Solar System and beyond, and for human colonization. In particular, the moon has been suggested as the ideal location for solar power generation (and subsequent microwave transmission to earth via satellite relay stations). A conference on Space Solar Power sponsored by NASA and NSF (and organized by USC faculty) included several papers on this topic (http://robotics.usc.edu/workshops/ssp2000/).

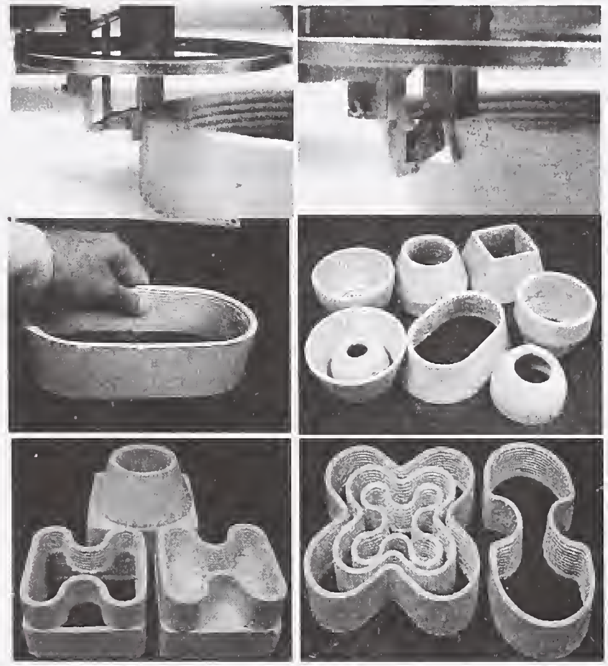


Figure 7. CC in operation and representative 2.5D and 3D shapes

Construction of solar collectors may be possible by using robots to assemble panels of photovoltaic cells shipped from earth. However, the cost of shipping these panels may be prohibitive. In fact, the materials needed to manufacture phovoltaic cells are all present in the lunar regolith, so that it may be more

Figure 8. Sections with cavities made with CC and filled with concrete (largest dimension in these figures is 10")

practical to build a "factory", perhaps by Contour Crafting technology, rather than shipping panels from earth. Furthermore, once solar power is available, it should be possible to adapt the current Contour Crafting technology to the lunar and other environments to use this power and in situ resources to build various forms of infrastructures such as roads and buildings. The lunar regolith, for example, may be used as the construction material. Other researchers have shown that lunar regolith can be sintered using microwave to produce construction materials such as bricks. We envision a Contour Crafting system that uses a limited amount of water to form a small batch of the lunar material into paste form. Once extruded and deposited in the desirable location, the water could be extracted and recycled for the next batch of material. Microwave sintering using solar power can be integrated into the system and hence progressively cured structures of various complexities could be built.

One of the ultimate goals of the Human Exploration and Development of Space (HEDS) program of NASA is colonization, i.e., building habitats for long term occupancy by humans. We believe that the Contour Crafting technology is the ideal method for such construction.

ACKNOWLEDGEMENT: This material is based upon work supported by the National Science Foundation under Grants No. 9522982 and 9615690, and by a grant from Office of Naval Research.

REFERENCES

Balaguer C., M. Abderrahim, S. Boudjabeur, P. Aromaa, K. Kahkonen, S. Slavenburg, D. Seward, T. Bock, R. Wing, & B. Atkin (2002), "FutureHome: An Integrated Construction Automation Approach," IEEE Robotics & Automation Magazine, pp 55-66.

Khoshnevis, B. (1998), "Innovative rapid prototyping process makes large sized, smooth surfaced complex shapes in a wide variety of materials", *Materials Technology*, Vol. 13, No. 2, pp. 52-63.

Khoshnevis, B., Russell, R., Kwon, H., & Bukkapatnam, S. (2001-a), "Contour Crafting – A Layered Fabrication Technique", *Special Issue of IEEE Robotics and Automation Magazine*, 8:3, pp 33-42.

Khoshnevis, B., Bukkapatnam, S., Kwon, H., & Saito, J. (2001-b), "Experimental Investigation of Contour Crafting using Ceramics Materials", *Rapid Prototyping J.*, Vol. 7, No.1, pp. 32-41.

Kulkarni, P., and Dutta, D. (1999), "Deposition Strategies and Resulting Part Stiffnesses in Fused Deposition Modeling", *Journal of Manufacturing Science and Engineering*, ASME, No.1, pp. 93-103, February.

Pegna, J. (1997), "Explorary investigation of solid freeform construction", *Automation in construction*, Vol. 5, No. 5, pp. 427-437.

Shrunk, D. Sharpe, B, Cooper, C., and Thangavelu, M., (1999), *The Moon: Resources, Future Developments, and Colonization,* Wiley-Praxis Series in Space Science and Tech., NY.

Warszawski, A., and Navon, R. (1998) "Implementation of robotics in buildings: current status and future prospects," *Journal of Construction Engineering and Management*, Vol.124, No.1, pp. 31-41.

Zak, G., Sela, M. N., Park, C. B., and Benhabib, B. (1999), "Layered-Manufacturing of Fiber-Reinforced Composites", *Journal of Manufacturing Science and Engineering*, ASME, Vol. 121, pp. 448-455, August.

Rapid Human-Assisted Creation of Bounding Models for Obstacle Avoidance in Construction

by

McLaughlin, J. [1], Haas, C. [2], Liapi, K. [3], Sreenivasan, S.V. [4], and Kwon, S. [5]

ABSTRACT: State-of-the-art construction equipment control technology creates the opportunity to implement automated and semi-automated object avoidance for improved safety and efficiency during operation; however, methods for constructing models of local objects or volumes in real-time are required. A practical, interactive method for doing so is described here. The method: (1) exploits a human operator's ability to quickly recognize significant objects or clusters of objects in a scene, (2) exploits the operator's ability to acquire sparse range point clouds of the objects quickly, and then (3) renders models, such as planes, boxes, and generalized convex hulls, to be displayed graphically as visual feedback during equipment operation and/or for making proximity calculations in an obstacle detection system. Experimental results indicate that bounding models can be created rapidly and with sufficient accuracy for obstacle avoidance with the aid of human intelligence and that human-assisted modeling can be beneficial for real-time construction equipment control.

KEYWORDS: construction automation, workspace modeling, bounding box, convex hull, obstacle avoidance, laser range finder

1. INTRODUCTION

Recent research indicates that several applications such as earth moving, heavy lifting, and material handling can benefit from the use of graphical models of equipment and workspace [1], [2], [3], [4]. Real-time interference checking for obstacle avoidance is also possible using local area graphical models. Laser range scanners are fast becoming popular tools for collection of three-dimensional range data for construction site modeling [5]. These methods can produce very detailed models of the scanned scene, which are useful for obtaining as-built drawings of existing structures, however the computational and data acquisition time burdens preclude the methods from being used on site for the real-time decision-making. Overall, modeling times for these laser range scanners are on the order of hours or days. The dynamics of a construction site require modeling times on the order of seconds or minutes.

The dynamic nature of the construction environment requires that a real-time local area modeling system be not only rapid but also capable of handling the changing and uncertain work environment. The approach taken in this research relies on a human's cognitive ability to recognize and classify

[1] Engineer, Los Alamos National Lab, NM, jtmclau@sandia.gov
[2] Professor, Dept. of Civil Engrg., The University of Texas, Austin, TX 78712, haas@mail.utexas.edu
[3] Assist. Professor, Dept. of Civil Engrg., The University of Texas, Austin, TX 78712, kliapi@mail.utexas.edu
[4] Assoc. Professor, Dept. of Mech. Engrg., The University of Texas, Austin, TX 78712, sv.sreeni@mail.utexas.edu
[5] Graduate Research Assistant, Dept. of Civil Engrg. The University of Texas, Austin, TX 78712, swkwon@mail.utexas.edu

objects in the workspace. Much research has been conducted on automatic object recognition for model generation, but these methods are neither robust nor efficient enough for real-time modeling in construction. The goal here is to balance human discernment and efficient range data acquisition with the proper exploitation of the computer in the areas of model generation, interference checking and avoidance control.

2. RAPID WORKSPACE MODELING

The following three sections describe three modeling methods that were developed and found to be useful for rapid workspace modeling for obstacle avoidance: 2.1) workspace partitioning, 2.2) convex hulls, and 2.3) tight-fitting bounding boxes. It should be emphasized that all of the above methods were developed for compliance in a local obstacle detection system. Since high numbers of objects in a workspace compounds the effects of slow distance computations, because pair-wise comparisons of all manipulator links to all objects must be made continuously, all the modeling methods described below take advantage of low numbers of range points for fast data acquisition and modeling as well as planar surfaces for quick proximity calculations.

2.1 Workspace Partitioning

The first and simplest model described is a finite plane (or infinitely thin wall) used for partitioning a workspace. Only three points in space are necessary to define a plane. However, a least-squares approach using more than three points is very useful to ensure that the plane is placed where the operator had intended it to be. The mathematics implemented are described in [6]. Floors, walls, and ceilings (i.e., rooms) can be quickly modeled this way by picking just a few points.

2.2 Convex Hulls

In three-dimensional space, the convex hull of a set of points is the smallest convex volume that contains the points. There are good reasons for using convex hulls for rapid obstacle avoidance modeling:

- Convex nature makes the hull inherently conservative
- Any number of points can be picked, anywhere
- The resulting hull consists of planar faces for fast distance computation

The algorithm used in this research is an incremental algorithm by Barber, Dobkin, and Huhdanpaa that successively adds a point to the convex hull that was generated by using the previously processed points [7]. The details of the algorithm are not discussed here for the sake of brevity.

2.3 Tight-Fitting Bounding Boxes

Much of the same benefits of convex hulls also hold for bounding boxes. Like generalized convex hulls, boxes are convex polyhedrons. Boxes are useful for acting as a simple outer shell that can hide a more detailed and precise model underneath. The primary reason for doing this, relative to obstacle avoidance, is so that at large distances, where manipulator movements are small compared to the overall distance from the manipulator base to the object, the manipulator's detection system is not forced to deal with a complex model. As the manipulator approaches the object, the object's details become more relevant, so the outer box is removed. This multi-layered modeling approach is useful in cluttered environments where high numbers of complex models would stifle an obstacle detection system. The algorithm developed to create the tight-fitting bounding box is described in [6].

3. EXPERIMENTS

Modeling experiments were conducted to determine the applicability of the modeling methods above. The actual mechanism used by the operator for the point collection and the interface issues therein were not the focus of this research. Rather, the human's ability to recognize the important features in a scene as well as the points needed to define models of prescribed geometry of those features was the focus. Twenty test subjects performed the modeling experiments. The experiments aimed to satisfy two sub-objectives as well. First, the relation between speed and accuracy was sought. While speed is obviously the driver, adequate accuracy is essential to obstacle avoidance and must not be abandoned for the sake of speed. Second, test subjects were asked to repeat certain tasks so that a learning curve could be observed.

3.1 Experimental Setup

Figure 2 is a picture of the mock scene that was set up in the construction automation laboratory for this and other modeling experiments. Referring to the figure, four models were used in the experiments described here:

1) A vertically constrained wall obtained by picking points on each of the three orange construction cones in the rear of the scene
2) Three convex hull/tight-fitting bounding box combination models of the wood box, pipe rack, and junk pile

The three construction cones were placed somewhat linearly so that the resulting vertical wall, as seen in the graphic display window by the test subjects, would unambiguously coincide with the cones. The width of the wall was arbitrarily forced to equal the distance between the two furthest cones and the height was arbitrarily set to six feet to be definitive. The wood box, pipe rack, and junk pile were each used for the convex hull and bounding box modeling. These three objects were chosen for their variations in geometry, complexity, size and the number of points required to define the convex hull. The junk pile was just a random assortment of pipes, boards, and a pick ax.

Data acquisition was accomplished using a laser range finder mounted on a two degree-of-freedom pan and tilt unit (PTU). The laser was directed via a trackball controller and the graphical models were displayed on the computer screen using the Matlab™ GUI (Figure 3). For details on the retrieval of the laser distances and pan and tilt angles as well as the forward kinematics of the system see [6].

3.2 Experimental Method

Prior to performing any of the modeling, each test subject was given some motivation by explaining the nature of the research project and rapid world modeling in general. They were asked to imagine themselves with the task of the equipment operator who needs to quickly create a graphic model of the workspace scene by picking various points on the objects using the PTU-mounted laser range finder. The operator as visual feedback would then use this graphic model during the manipulation task as well as by the obstacle detection system. Next, the test subject was introduced to the data acquisition system (Figure 3). Once the subject felt comfortable with the system, modeling began. Each model was displayed graphically using the Matlab™ GUI immediately after it was modeled so that the experimenter could see the results and the effects of the decisions that were made. The graphic workspace model was updated with each new model so that by the end of the last object model, the experimenter had a complete local graphical workspace model of the scene.

The time was recorded for each modeling exercise and commenced on the registration of the first distance measurement of the laser and ended on the registration of the last distance measurement. Qualitative observations were made and recorded as each test subject used the system. A means of quantifying the accuracy or conservativeness with which a test subject could model an object or objects by picking points to create a convex hull was also necessary. This was accomplished by developing a ray-tracing algorithm. This algorithm essentially compares the smallest convex volume that could encompass an object with the convex hull created by the test subject. It is detailed in [6].

3.3 Scoring Function

In addition to the ray-tracing algorithm, which enabled quantification of the convex hull modeling accuracy, a means of quantifying the overall convex hull modeling performance of each of the test subjects was necessary. Four related criteria emerged as the most significant in determining the effectiveness of a convex hull modeler:

1) Accuracy - as discussed above

2) Time - total elapsed time acquiring points per object
3) Efficiency - the number of convex hull points versus the total number of range points per object
4) Number of Missed Points - the number of missed points as detected by the ray-tracing algorithm

A scoring function was formulated that combines these factors and is detailed in [6].

3.4 Experiment Results

Referring to the picture of the workspace scene in Figure 2, Figure 4 is an example of the completely modeled scene done by one of the test subjects of the first group. Notice that each of the objects has been modeled with the appropriate method (wall - planar fit, wood box/pipe rack/junk pile - convex hull/bounding box). Modeling the pipe as a cylinder is described in [2], [6] and is not discussed here due to length restrictions.

The learning curve was not monotonic for about half of the subjects. In fact, it was observed in most of these cases that as the subject's understanding of the convex hull modeling approach grew stronger and enthusiasm for performing the experiment diminished, the test subject would attempt to model the object with a minimum number of points. This led to some missed points and lower accuracies, which despite an improved time, resulted in a lower score. This is apparent in Table 1, which shows the averages for each of the four metrics and the resulting score for both attempts of the second group. Notice that the average number of missed points for the second attempt at modeling the junk pile was actually higher than the first attempt, despite an overall improvement in score. The improvement in score was most dramatic for the pipe rack, which makes sense since it was modeled first in the sequence.

The most significant result, as shown in Table 1, is the average deviations of the experimenter's convex hulls from the control hulls. These average deviations were roughly an inch after two attempts for both the pipe rack and junk pile. Moreover, the median deviations were even smaller than the averages (0.92" for the pipe rack and 0.80" for the junk pile). Deviations this small are quite negligible with respect to large construction manipulators where the closest allowable distance from the manipulator to an obstacle would be larger.

4. OBSTACLE AVOIDANCE SIMULATION

An obstacle avoidance simulation was performed to demonstrate the applicability of the modeling methods to obstacle detection for the purposes of equipment operator feedback and control. Since construction equipment tends to be large and massive, inertia is an extremely important factor to be monitored for safe navigation. Thus, the simulation was designed to monitor manipulator link velocities as well as positions. The simulation consisted of a three-dimensional, three degree-of-freedom robot traversing over a box. Initial and final joint angles and a total elapsed time were specified. The joint paths were then forced to follow smooth fifth-order curves. The Gilbert, Johnson, and Keerthi algorithm for computing the minimum distance between convex polyhedra in three-dimensional space was used as a fast method of proximity calculation [8]. The velocity was accounted for by running a forward dynamic sub-simulation at each control time step (100 Hz) to see where the manipulator would stop given its current joint angles and velocities as initial conditions. The actual positions as well as the projected positions from the sub-simulation were put into an artificial potential function as feedback output [6]. The simulation indicated that obstacle avoidance for a construction manipulator instrumented with feedback control would be feasible in real-time given the relatively simple models described in this paper.

5. CONCLUSIONS

Three modeling methods were found to be useful for construction site modeling: workspace partitioning, convex hulls, and bounding boxes. The low deviation values (about one inch), and the low modeling times (about 2-3 minutes) in Table 1 indicate that a human-guided laser range finder can model construction site objects significantly faster than current methods and with sufficient accuracy. In contrast to an autonomous scanner, the human can quickly recognize the important features of the scene and then direct the laser accordingly, decreasing data acquisition time and, consequently, computational time due to the lower number of points.

6. REFERENCES

[1] Stentz, A., Bares, J., Singh, S., and Rowe, P., "A Robotic Excavator for Autonomous Truck Loading," Autonomous Robots, Vol. 7, No. 2, pp.175-186, September 1999

[2] Cho, Y. (2000), "Human-Assisted Rapid Workspace Modeling for Construction Equipment Operations," PhD Dissertation, The University of Texas at Austin

[3] Kim, Y.S. and Haas, C. (2000) "A Model for Automation of Infrastructure Maintenance using Representational Forms," Journal of Automation in Construction, Vol.10, No.1.

[4] LeBlond, D., Owen, F., Gibson, G., Haas, C., and Traver, A. (1998). "Performance Testing and Characterization of Advanced Construction Equipment," ASCE Journal of Construction Engineering and Management, Vol. 124, No.4, pp. 289-296.

[5] Cheok, G. S., Lipman, R. R., Witzgall, C., Bernal, J., and Stone, W. C. (2000), "Field Demonstration of Laser Scanning for Excavation Measurement," Automation and Robotics in Construction XVII

[6] McLaughlin, J. T. (2002), "Rapid Human-Assisted Creation of Bounding Models for Obstacle Avoidance in Construction," Masters Thesis, The University of Texas at Austin

[7] Barber, C., Dobkin, D., Huhdanpaa, H. (1996), "The Quickhull Algorithm for Convex Hulls," ACM Transactions on Mathematical Software, Vol. 22, No. 4, pp. 469-483, December 1996

[8] Gilbert, E. G., Johnson, D. W., Keerthi, S. S. (1998), "A Fast Procedure for Computing the Distance between Complex Objects in Three-Dimensional Space," IEEE Journal of Robotics and Automation, Vol. 4, No. 2

ACKNOWLEDGEMENTS

This paper is based on the research funded by the National Science Foundation (Grant#: CMS-0000137) and the National Institute Standard and Technology (Project NBR: NA1341-02-W-0742). The authors gratefully acknowledge their financial support and encouragement throughout this study.. The authors gratefully acknowledge their financial support and encouragement throughout this study.

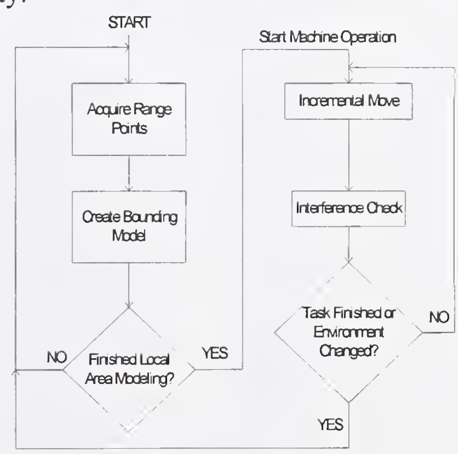


Figure 1. Overall Construction Equipment Operation Modeling Process

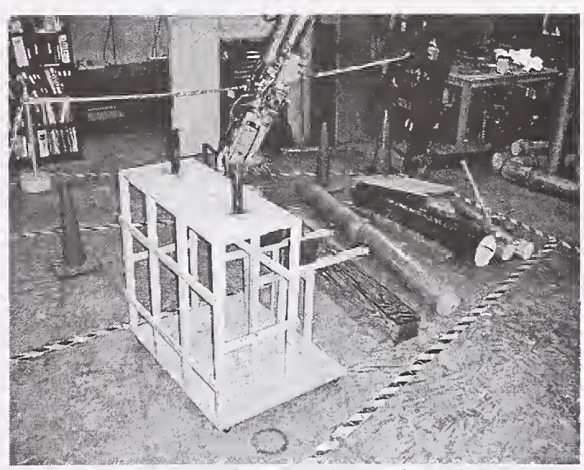


Figure 2. The Scene for Experimental Modeling

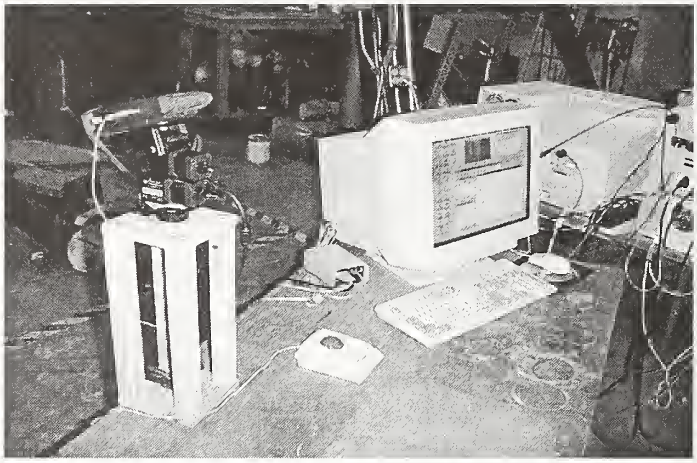


Figure 3. Laser, PTU, Trackball Control, and Data Acquisition Software Interface

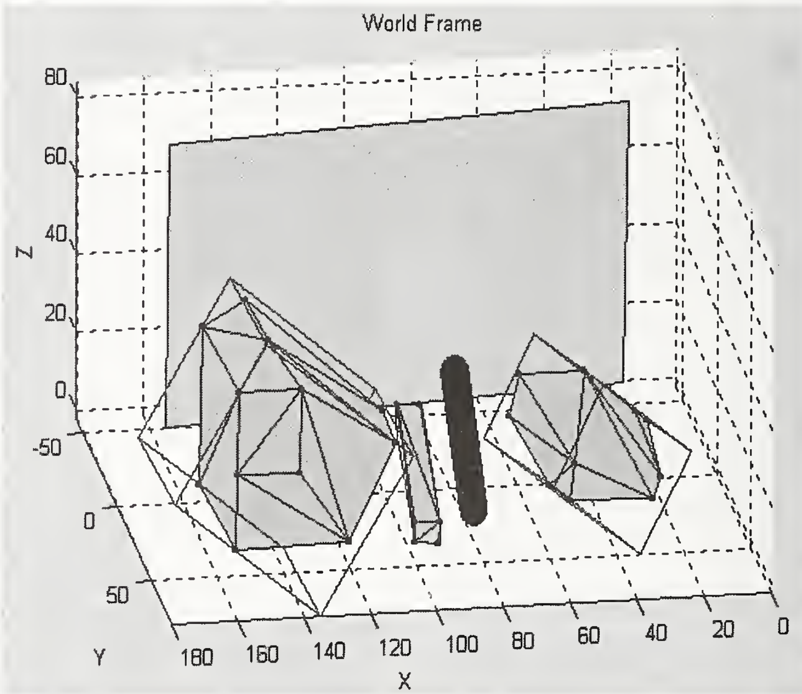


Figure 4. Graphic Model of Workspace Scene (dimensions are inches)

Table 1. Average Values for the First and Second Attempt at Modeling the Pipe Rack and Junk Pile

	Pipe Rack		Junk Pile	
Averages	**1st**	**2nd**	**1st**	**2nd**
Deviation (in)	2.42	1.16	1.12	0.97
Time (min)	5:51	2:33	3:30	2:17
# Range Points	26.3	14.6	16.6	13.1
# Hull Points	16.0	11.9	13.3	11.6
H/R Ratio	0.68	0.86	0.86	0.90
# Misses	1.0	0.1	0.4	0.7
Score	**39**	**76**	**72**	**76**

COMPUTER TECHNOLOGIES IN CONSTRUCTION ROBOTS CONTROL

Thomas Bock*, Dimitry Parshin, Rudolf Neudorf**, Alexej Bulgakow***

**Technical University Munich, bulgakow@bri.arch.tu-muenchen.de*
*** Rostov State Building University, RUSSIA, pragma_plus@mail.ru*

Abstract: Construction is of great interest for applying practical and theoretical results in the field of robotics and computer technologies. The paper presents the composition and special features of kinematics and dynamic models of construction robots, control algorithms synthesis in terms of synergetic approach, planning of robot's movements based on computer techniques. Recommendations on computer technologies application, construction robots control, control programmes development and operators teaching are given.

Keywords: automated control system, construction, computer technologies, kinematics, dynamic, model, robot.

1. INTRODUCTION

Construction robotics is of great interest for practical application of the results while performing automation of mounting, finishing, placement of concrete and other kinds of work. However robots application in conditions of construction sites is connected with technological, operational and some other characteristic properties. Successful solution of the tasks dealing with construction operations robotization requires not only the development of special-purpose manipulators but the application of new approaches while solving such problems as control, management and preparation of control programs. In connection with the mentioned above we suggest to use computer technologies in manipulation system control, trajectory path planning, developing routes for handling the parts, developing control programs in the teaching mode, robots analytic programming, robots task-level programming and operators teaching for construction robots taking into account their specific features.

2. MATHEMATICAL MODELS FOR ROBOTS

In the basis of a computer control technology lie mathematical techniques representing the set of models of different levels: from the models of decision making to particular models for carrying out controlled movements of some degrees of freedom. Control algorithms are constructed on the basis of the obtained kinematics and dynamic models. Kinematic models include equations determining position and speeds of the system elements and specifying transport, installation, transfer and orienting movements of construction robots. The position of coordinate systems while developing kinematic model must ensure conventional representation of mechanisms location and simplify the process of coordinates transformation. For self-mobile construction robots the basic coordinate system $X_mY_mZ_m$ should be connected with the manipulator rotation axis and the axis X_m is directed along the self-propelled truck. (fig.1). For the rotational degrees of freedom, axes

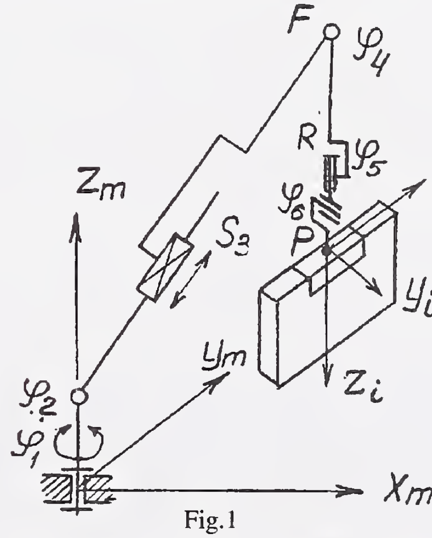


Fig.1

of which are perpendicular to the links (2nd, 4th, 6th pairs) it is convenient to combine axes X_i with rotation axes and axes Z to be led along the link. For the translational degree of freedom and the rotational one with the location of the axis along the link it is necessary to direct axes Z_j along the axes of movement and axes X_j to combine with axes X_{j-1} of the previous degrees of freedom. Direction of coordinate system axes $X_iY_iZ_i$ of a gripper or a technological tool is defined by a robot. For mounting robots it is convenient to direct axis Z_i

along the end link and axis X_i - along the traverse of the gripping device. In finishing robots we recommend to direct axis Y_i along the end link and axis Z_i to the plane of the mechanism. In this case it is expedient to construct kinematics model on the basis of position T_R and orientation T_{OR}. Using these matrices we candefine the position and orientation of the tool: $T_{\rm p} = T_{\rm F} \cdot T_{\rm or}$, and on the basis of this matrix we get the system of equations to determine coordinates of point P in the coordinate system of the robot $X_mY_mZ_m$:

$$X_P = l_6 \cdot c\varphi_1 s\varphi_5 s\varphi_6 + l_6 \cdot s\varphi_1 c(\varphi_2+\varphi_4)c\cdot \varphi_5 s\varphi_6 + (l_5 + l_6 c\varphi_6)s\varphi_1 s(\varphi_2+\varphi_4) + (S_3 s\varphi_2 + l_4 s(\varphi_2+\varphi_4))s\varphi_4;$$
$$Y_P = l_6 \cdot s\varphi_1 s\varphi_2 s\varphi_6 + l_6 \cdot c\varphi_1 c(\varphi_2+\varphi_4)c\cdot \varphi_5 s\varphi_6 + (l_5 + l_6 s\varphi_6)c\varphi_1 s(\varphi_2+\varphi_4) + (S_3 s\varphi_2 + l_4 s(\varphi_2+\varphi_4))c\varphi_1;$$
$$Z_P = l_1 + S_3 c\varphi_2 + l_4 \cdot c(\varphi_2+\varphi_4) + (l_5 + l_6 c\varphi_6) c(\varphi_2+\varphi_4) + l_6 \cdot c\varphi_5 s\varphi_6 s(\varphi_2+\varphi_4).$$

where $C\varphi_i = \cos(\varphi_i)$, $S\varphi_i = \sin(\varphi_i)$.

In general manipulator kinematics is represented as a system of equations being the functions of generalized coordinates:

$$x_{\rm p} = f_1(q_1, q_2, ..., q_6);$$
$$y_{\rm p} = f_2(q_1, q_2, ..., q_6);$$
$$z_{\rm p} = f_3(q_1, q_2, ..., q_6);$$
$$\theta_{\rm p} = f_4(q_1, q_2, ..., q_6);$$
$$\beta_{\rm p} = f_5(q_1, q_2, ..., q_6);$$
$$\alpha_{\rm p} = f_6(q_1, q_2, ..., q_6).$$

To solve inverse kinematics problems we recommend to use approximate relations based on iteration methods. More exact values of generalized coordinate for a construction robot $q_i^{(k)}$ on k-th step of the iteration cycle are determined through approximate values $q_i^{(k-1)}$ on k-1 step:

$$q_i^{(k)} = q_i^{(k-1)} + U_i^{-1}\left[(T_i - T_i^{(k-1)}) + \sum U_{ij} q_j^{(k-1)}\right],$$

where $U_{ij} = \dfrac{\partial T_i}{\partial q_j}$ is a partial derivative of transformation matrix T_i along the j-th generalized coordinate.

When controlling the robot's movement the data about the speeds of the grip motions and the point defining the transport path of the movement are necessary. At the given speeds of the manipulator links motions the projections of the tool linear speed are defined by the equations:

$$v_x^{(\rm P)} = \sum_{i=1}^{4} \frac{\partial x_p}{\partial \varphi_i}\dot{\varphi}_i + \frac{\partial x_p}{\partial S_3}\dot{S}_3;$$
$$v_y^{(\rm P)} = \sum_{i=1}^{4} \frac{\partial y_p}{\partial \varphi_i}\dot{\varphi}_i + \frac{\partial y_p}{\partial S_3}\dot{S}_3;$$
$$v_z^{(\rm P)} = \sum_{i=1}^{4} \frac{\partial z_p}{\partial \varphi_i}\dot{\varphi}_i + \frac{\partial z_p}{\partial S_3}\dot{S}_3.$$

To simplify the development of kinematic models we suggest to use the method structure decomposition, the essence of which is in breaking the architecture into elementary type structures. This allows to have transformation matrices for each type module that connect its input and output values.

One of the effective methods of modelling dynamic characteristics of robot's effectors is dynamic models construction. According to them we select control algorithms, form optimal laws of control and determine control forces in degrees of freedom. While developing mathematical model for a robot there appears a question about the choice of mathematical device. For construction robots it is convenient to apply motion equations in Lagrangian form. In this case manipulator dynamics is described by a set of interrelated non-linear differential second-order equations of the form:

$$\begin{cases} \sum_{i=1}^{6} a_{1i}\ddot{q}_i + \sum_{i=1}^{6} b_{1i}\dot{q}_i + \sum_{i=1}^{5}\sum_{j=i+1}^{6} C_{1ij}\dot{q}_i\dot{q}_j + G_1 = M_1^{\Sigma} \\ \vdots \\ \sum_{i=1}^{6} a_{6i}\ddot{q}_i + \sum_{i=1}^{6} b_{6i}\dot{q}_i + \sum_{i=1}^{5}\sum_{j=i+1}^{6} C_{6ij}\dot{q}_i\dot{q}_j + G_6 = M_6^{\Sigma} \end{cases}$$

where $a_{ki}(m_i, l_i, q_i)$ - functions characterizing centropetal forces; $b_{ki}(m_i, l_i, q_i)$ and $c_{kij}(m_i, l_i, q_i)$ - functions depending on generalized coordinates and characterizing Coriolis and centrifugal forces; $G_i(q_1, q_2, ..., q_6)$ - potential forces.

Summary moments acting in degrees of freedom are generally determined by the sum

$$M_k^{\Sigma} = M_d^{(k)} - M_f^{(k)} - M_l^{(k)},$$

where $M_d^{(k)}$ – moments developed by a drive; $M_f^{(k)}$ – movements of friction degrees of freedom; $M_l^{(k)}$ – movements of external forces. The moment of a drive $M_d^{(k)}$ are defined by the moment of a motor $M_m^{(k)}$, efficiency $\eta^{(k)}$ and transmission ratio of the drive:

$$M_d^{(k)} = [M_m^{(k)} - J_m^{(k)} \frac{d\omega_k}{dt}] \cdot i_r^{(k)} \eta_d^{(k)},$$

where $J_m^{(k)}$ – moment of inertia of movable parts connected with the motor shaft; ω_k – angular speed. The load moment $M_l^{(k)}$ for each drive is defined by the weight of the working tool and depends on links position:

$$M_l^{(k)} = f(m_l, q_k, q_{k-1}, q_6),$$

A dynamic model developed on the basis of the described equations is added up by the control equations:

$$T_m^{(k)} \frac{dM_m^{(k)}}{dt} + M_m^{(k)} = k_b^{(k)} U_b^{(k)} - k_e^{(k)} \omega_k,$$

where $T_m^{(k)}$ – time constant of a motor anchor chain; $k_b^{(k)}, k_e^{(k)}$ – coefficients of transformation and self-induction.

3. ROBOTS CONTROL ALGORITHMS

They include the set of algorithms of motions planning and formation of controls for manipulator degrees of freedom. In performing this the parameters correction of the manipulator motions is carried out discretely in time interval Δ(t) and the planed trajectory and he laws of changing generalized coordinates provide continuity of the functions, their first and second derivatives. The new values of the motion parameters are generated from he periods of 1-10 ms. An effective principle of developing control algorithms for robots with complex dynamics is a synergetic approach. It is based on separation of motions in complex dynamic systems and approximation of the main properties of the controlled object with a simplified mathematical model. Application of these models with significant non-linearity of the object is limited by states space and time interval. At each interval linearization is permitted and asymptotically stable movement of the system is provided. Such an interval-approximation control possesses robustness as at each control interval adaptation to the properties of the object changing in time and space is ensured. To perform interval-approximation control a dynamic model of a robot as a controlled object is represented by equations of the form

$$\dot{\overline{x}}(t) = F(t, \overline{x}(t), \overline{u}(t)),$$

where F is an allowing approximation vector-function; argument of which are vectors of state $\overline{x}(t)$ and control actions $\overline{u}(t)$. At about point $\overline{x}_{i0}$ and $\overline{u}_{i0}$ of an integrated space of state and input actions we find an approximating dependence

$$F_{ia}(t, \overline{x}_{ia}, \overline{u}(t)),$$

where $\overline{x}_{ia}$ an evaluation vector of the object state with an approximating model in the *i*-th area of state under consideration. For the interval at the moment t_i we determine χ_i – a vicinity of initial conditions of a state vector $x_{i0} = x(t_i)$ and ν_i – a vicinity of current value of a control vector $u_{i0} = u_e(t_i)$, when evaluation error of its own and forced variations of state $\overline{x}$ with of approximation model

$$\overline{x}_{ia}(t) = F_{ia}(i, x_{ia}(\mathrm{t}), \overline{u}(t))$$

does not exceed some admissible value

$$\|\overline{\varepsilon}_{ix}(t)\| = \|\overline{x}_{ia}(t) - \overline{x}(t)\| \leq \varepsilon_m.$$

This allows to find the values of maximal admissible time interval $\tau_i \in [t_i, t_{i+1}]$ at which the legitimacy of approximation is provided. In this case at $\tau_i \leq \tau_{\max}$ and at exact estimation of the object initial state the control law is calculated according to an approximating model: $\overline{u}(\overline{x}, t) \approx \overline{u}(i, \overline{x}_{ia}, t)$.

Control at an attributed interval τ_i providing stable system movement with the given quality is performed in terms of the methods of optimal model control or methods of the system synthesis according to the desired characteristics. Synthesized control laws $\overline{u}(i, \overline{x}_{ia}, t)$ at the whole temporal control range represent a piece function implementing the algorithm for forming control law with variable parameters. The application of synergetic approach to the motion control of construction robots made it possible to obtain algorithms providing the solution of the control tasks with minimum of calculation operations. This allows to cut down the time for formulating the next vector of control actions, to reduce the interval of renewing control signals and to increase the accuracy of control. The algorithms applied make it possible to form control action according to the given differential equations describing the object to be controlled and the standard transitional processes.

The process of control algorithms development and control program preparation for construction robots is greatly simplified when computer technology is used. Applying methods of mathematical modeling, computer graphics we can perform the search for optimal trajectories of movements, carry out simulation of the processes of tool motion, provide division of trajectories into sections and form data base of ponts of control. The robot motions planning is based on non linear control algorithms and carried out with the account of limitations both for the trajectory itself and the manipulator links displacement. Correction of movements parameters is carried out discretely at time interval $\Delta(t)$, and the planned trajectory provides continuity of the function. Values of movement parameters are generated at a period of 1-5 ms. As a result of robot motions planning the obtained laws of changing internal coordinates $q_k[nT]$ are then transformed into control actions $U_k[nT]$ providing motion stability relative to the chosen programmed trajectory. Incorporation of the events processor into the system makes it possible to synchronize the robot's operation with other technological equipment that is especially important for construction-mounting operations.

4. PARAMETERS, CONTROL, PROGRAMS PREPARATION

Construction-mounting operations require constant control of the equipment condition and the technological process parameters. The application of computer technologies allows to obtain information from the sensors about environment situations, means of control concerning the objects positions, to process input data, to display the data by the operator's request at the pre-set period. Visual control of technological operations performance in real time allows to simplify to the utmost the operator's work. Construction – mounting robots should have programs to control the main parameters of movements: the position of each degree of freedom, speed of links movement, gripper (tool) coordinates and speeds of its motion. According to the data read out of position sensors and tachometer generators of the manipulator degrees of freedom tool actual coordinates and its orientation are calculated as well as a vector of tool motions speeds. The calculations are fulfilled on the basis of the algorithms of kinematic model. Meanwhile by operator's request positioning errors can be formed and displayed graphically in the function of time or trajectory.

An important trend in computer technologies application in construction robotics is preparation of control programs which can be carried out outside the construction site on the model of a robot or robotic system. The application of information technologies makes it possible to create a robot's virtual model and a cross-system for preparing and testing control programs and their translation into the codes of the robot control system. A virtual robot which is displayed in 3-D graphics and moves in real time makes up the basis of such system for developing programs. The developed algorithms and models are put into the basis of the integrated system including the subsystem of 3-D modeling, the system of programming and separate modules for carrying out research and educational tasks. The package of algorithms and programs allows to calculate automatically the positions of all robot's joints participating in performing the pre-set movements, to form control actions and to test their performance on the robot's model. In the process of fulfilling the movements the violation of the introduced limitations for the gripper motion may occur. The characteristic feature of the program developments system is the opportunity of testing control algorithms in different modes with the convenience of imitating the processes of collecting data and control. The described models and algorithms make up the basis of an integrated system of software for civil engineering works. It has a multiwindow interface with functional, graphic and table windows, which can be adjusted to the given kinematic structure. Incorporation of 3-D objects info the user interface makes it possible to simulate visually the operations performed by the robot, to teach the operations.

5. CONCLUSION

The presented material is the result of developments in the field of automation and robotization of civil engineering works. On the basis of the described approaches kinematic and dynamic models construction-mounting and finishing robots have been developed, control algorithms have been synthesized according to synergetic approach to movements control. A software integrated system is of practical interest, it allows to control parameters, visualize robot's work, task in a graphical or table form of movement trajectories, it also provides preparation of programs and operators teaching.

6. REFERENCES

1. V. Zagorodnjuc, D.Parchin, The Construction Robotica (Mockow,1990).

EQUIPMENT OPERATOR TRAINING IN THE AGE OF INTERNET2

Leonhard E. Bernold, *Associate Professor*
Justin Lloyd, *RA*
Mladen Vouk, *Professor*
Construction Automation & Robotics Laboratory, North Carolina State University, Raleigh, NC, 27606

Abstract: Considering the established needs for operator training in the construction industry and vastly improved data communication using Internet2, the feasibility of an effective Internet based training system for backhoe operators is increasing. This paper presents ongoing work on a prototype system designed with the idea that eventually any electronically controlled backhoe can be adapted and connected to the Internet. Experimental tests have shown that 3-D, real time video in combination with force and sound feedback provide a data-rich control environment for the trainee. This paper also discusses how the availability of force and spatial position data can be used for running an online training program that provides the user with a series of structured tutorials and practice sessions. System features of a Virtual Coach might include the haptic feedback featured in the prototype system, depending on its demonstrated relevance to training efficacy, as well as digitized verbal information, including instructions, commands and real-time advice.

Keywords: training, skill assessment, information technology, motion control, construction equipment

Introduction

The most common task of backhoe operators is to cut soil with a bucket and dump it on to a pile or into a truck. Skilled operators dexterously control the arm and bucket using two joysticks, as shown in Fig. 1, by monitoring the backhoe motion and the state of the process. The skill itself, however, is difficult to assess and even skilled operators can't describe their own "skillfulness".

Efficient on-the-job operator training is not only costly but is often not possible, requiring specialized equipment and an on-the-job trainer. However, only appropriate and extensive training enables operators to control large equipment safely and efficiently. Similar to other industries, construction tries to take advantage of computer simulators and virtual reality as training vehicles. An apparent problem with such tools is the lack of realism, which is very costly to create. In addition, research has shown that the key to acquiring the necessary motor skills to control complex systems, such as a backhoe excavator, is hands-on and coached training. (Cuqlock-Knopp et al., 1991)

The advent of Internet2 opens the opportunity to revolutionize training of large equipment operators. The large bandwidth and high-speed data transfer allows large amounts of data to be "pumped" through the Internet. In particular, live video can visually "connect" an Internet user with a remote site in near real time. In order to study the effectiveness of a distance training & learning concept, an Internet-Based Backhoe Operator Trainer (IBOT) was built and tested using the Internet as well as Internet2.

What is Skill?

How does one assess the skill of a backhoe operator today? The commonly heard answer is: "By watching him operate the equipment or measuring the time it takes him to load a truck." It becomes apparent that these skill measures are rather crude considering some common definitions of skill or skilled performance.

Paul Fitts (1964), a prominent motor skill investigator before his untimely death, listed four critical characteristics of skill: 1) goal directed, 2) exhibiting highly integrated and organized behavior, 3) acquired through practice and training, and 4) applied with ease. Fitts' definition of skill will be used as the basis for establishing performance measures that could be used to differentiate between an excellent and a not-so-perfect backhoe operator.

Measures of Operator Performance

Sage (1984) provides a definition of skilled performance that can be utilized to assess a backhoe operator. He characterizes a skilled performer as one who "can produce an output of high quality (such as fast or accurate) with a good deal of consistency. Skilled performance is also characterized by an appearance of ease, a smoothness of movement, an anticipation of variations in the stimulus situation before they arrive, and an ability to cope with these and other disturbances without disrupting the performance. Indeed, increasing skill involves a widening of the range of possible disturbances that can be coped with without disturbing the performance." Considering Sage's elaboration, we can establish the following list of measures for performance measurement: 1) quality of output, 2) consistency of operation, 3) smoothness of bucket motion, 4) proactive thinking, 5) adaptability to changing environments, and 6) capability for skill improvement. It is easy to recognize that these six traits are really based on earlier work by Fitts, who already had established a useful framework for skill acquisition earlier.

An Internet-Based Backhoe Operator Trainer

An Internet-based training system needs to provide the trainee a means to operate equipment without being situated right next to it, thus losing the ability to use his inherent senses as feedback channels. Figure 2 shows the architecture of the first prototype IBOT developed by the Construction Automation & Robotics Laboratory (CARL). The main components of the IBOT consist of: a) computer-integrated-backhoe (CIB), b) one PC running video communication software c) one computer for digital data communication, and d) human-computer interface at a remote location with joystick, audio, video, and data display. NetMeeting was found convenient for demonstrations but, because it is already old technology, is not sufficiently dependable to allow experimental tests with IBOT. A commercial force-feedback video game joystick is used to control the excavator. The remote control software utilized a combination of Windows sockets, Windows GUI with the WIN32 API and multi-threaded programming. Here, a client and a server program work together to convert un-calibrated joystick signals originating from any computer with a Windows 98 operating system and standard joystick port into a calibrated, useful signal output from the analog digital (A/D) conversion board in the control computer. The client and server programs both engage independently operating communication threads which run in continuous loops, where the server thread echoes all data received from the client to maintain synchronization between the programs, as well as to provide a channel for feedback data, if necessary.

Figure 3 a) shows the programmable game joystick with which the entire excavator can be operated, while Figure 3 b) presents a view of the operator control seat where the operator is replaced by a video camera. Two large soil boxes serve as mobile containers for digging and dumping. Overhead cranes are able to pick them up and move them if desired. This allows the rapid change from one type of soil to another.

Electronic Skill Measurement

The modeling of the joystick motions required to move an excavator bucket through soil is a complicated problem, in as much as there exist many ways of creating the same path. In the same vein, the forces that have to be created to cut through the soil not only depend on the strength and density of the soil, but also on the digging motion, which includes factors like velocity, acceleration and smoothness of the bucket motion. Thus, it seems reasonable to expect that the forces that are created during the operation could be used as one important measure of operator skill.

The learning of motor skills is special in that it concerns itself with the process that underlies movement. Of critical importance is the fact that motor skills cannot be measured directly, but only inferred by observing behavior or performance. However, using performance alone as the measure of skill is not sufficient since other factors such as motivation, fatigue, boredom, noise, and temperature also impact performance. Anybody who has worked with equipment operators can easily see that the speed of skill acquisition depends upon a variety of conditions, with practice being the most important of them. In order to eliminate the extraneous effects as much as possible, trainees were assessed continuously, an act which provided data that could be plotted. The resulting graphs are called learning curves or performance curves.

Preliminary experiments, where two test persons were asked to perform an identical digging task, have shown that there is a definite relationship between skill and sensory feedback from the computer integrated backhoe. Figure 4 presents the hypothesis that smoothness of motion, which is related to skill, can be translated into smoothness of force. As shown, one should expect that a highly skilled operator creates a fairly consistent force pattern along the entire task while a poorly skilled operator produces spikes.

In order to test the basic hypothesis, a set of comparative experiments was executed. Two distinctively different operators, an "Expert" (E) and an "Apprentice" (A), were given the task to dig in the soil box under similar conditions.

Identification of Primary Skill Indicators

The tests conducted were based on the digging scheme presented in Figure 4. In a first effort, the research team attempted to verify the assertion that there are, in fact, unique force characteristics in each of the three cylinders: 1) boom, 2) stick, and 3) bucket. Fig. 5 presents two force data sets collected during one dig executed by two test subjects.

Force Pattern

Reviewing the graphs, several differences can be easily identified. For example E is filling the bucket in approx. 11 seconds while A needs almost 18 seconds to do the same. The above hypothesis seems to hold by the force data. E's maximum stick force during digging is 2,700 kN, while A needs 7,600 kN to move the stick through the soil. In addition, the "Expert", exhibits a much smoother and more uniform force distribution when compared to the "Apprentice."

Spatial Pattern

A second set of sensory data output that was analyzed resulted from the three angle encoders mounted on the three joints of the arm. Figure 6 depicts the result of a kinematic transformation, according to the Denavit-Hartenberg notation, converting the angles into Cartesian coordinates representing the path of the bucket.

While the path coordinates generated by E did fit a cubic polynomial path function, A's path shape only fits a cubic polynomial with a rather large deviation error. This correlates well with the stick-force data, since it also showed a comparatively "rough" pattern.

Coached Backhoe Training

The important first job of a coach is to instill in the trainee's memory an appropriate idea of how to execute a task. The two most commonly used techniques are verbal instruction and demonstration. A good coach also motivates and reinforces learned skills. Furthermore, coached training should be adaptive, in that the demands of the training environment and the complexity of the task "grows" with the skill level of the trainee. Thus, a coached training program is individualized and constantly monitored by the coach to assess progress. This personalized "looking over the shoulder" method of training is hardly possible on an actual excavator. A virtual trainer that is able to read and interpret feedback data representing the performance of the trainee may serve the role of an actual coach.

The module of a telepresence excavator trainer may be modeled as a closed-loop system. It is able to read measures representing the trainee's performance as input to dependent output coefficients, such as the difficulty of the training task. Figure 7 presents the framework of such a closed-loop training system for backhoe operators.

Towards a Virtual Coach

Considering the established needs for operator training in the construction industry and the versatile control and performance evaluation characteristics of the IBOT, the development of a "Virtual Coach" training system for backhoe operators using the IBOT system is a natural goal. System features of the Virtual Coach might include the haptic feedback featured in the IBOT, depending on its demonstrated relevance to training efficacy, as well as digitized verbal information, including instructions, commands and real-time advice.

The primary advantage of a teleoperation training system over a virtual reality training system is the guaranteed realism it provides. Complex and unpredictable processes, instead of being mathematically modeled and simplified, actually occur and present the trainee with stimuli commensurate with real-world equipment operation. Indeed, this type of "virtual" instruction would provide the trainee with an experience nearly identical to an on-site training program with a human instructor, saving the money and danger of such training practices and simultaneously providing more extensive performance evaluation capabilities.

Studies indicate the reliable transfer of skills acquired through virtual training or telepresence to actual on-the-job tasks (Rose et al., 2000), provided a high similarity between the virtual and real environments. Employing real-time, high-resolution stereoscopic visual feedback, audio feedback, and joystick control, the IBOT system generates the same intrinsic feedback provided by an on-site backhoe training session, and creates a teleprescence experience practically identical to an actual backhoe system in operation. Thus, provided an appropriate program design, the IBOT system would undoubtedly enable the cost-effective remote training of backhoe operators.

Because of the range of extra evaluative tools and feedback features the IBOT offers, the integration of automated training and Virtual Coaching aspects into the IBOT system and software presents a complex design problem, requiring particular knowledge in the subjects of motor learning, motor control and backhoe operator skill. System features as straightforward as performance feedback, user instructions and feedback augmentation would require careful planning and design, taking into account the current knowledge base on the efficacy of different types of supplemental aid to motor learning. Motor learning research (Cuqlock-Knopp et al., 1991) suggests a number of subtle but consistent dependencies of long-term motor skill retention on knowledge of results (KR) scheduling, performance feedback structure, augmented feedback usage, task instructions and attentional focus. The Virtual Coach system will enable precise software control over each of these aspects of motor skill instruction, providing a means to not only train operators, but to actually evaluate and refine the techniques of the trainee

Summary

Retrofitting existing equipment with ruggedized sensors and data-storage devices has enabled the implementation of various schemes for improving equipment life and performance. By installing computer-integrated controls, a path for operating them tele-robotically has been created. This paper discusses still another opportunity for utilizing such an electronic infrastructure, namely electronic skill assessment and training. One measure of good operator performance is the smoothness of motion. Using experimental data it is shown that there are, in fact, distinctive differences between the force-pattern generated by an “Expert” and an “Apprentice” backhoe operator. The extensive amount of data that was collected using force sensors and angle-encoders allowed the identification of unique features not only of forces but of path and energy consumption as well. It is argued that it is possible to develop a feature-based “pattern-language” enabling the automatic characterization of individual skill levels.

On-the job training with expensive equipment is not an economical approach to training, since it increases the chances of damage to the equipment, possible accidents, and a drastically reduced production. This paper offers a new approach to the dilemma of cost effective hands-on training. The discussed alternative takes advantage of the Internet as a communication tool that allows a trainee to practice with a stationary backhoe from his home. Named the Internet-Based Backhoe Operator Trainer (IBOT), the system has been tested using the Internet as well as Internet2.

References

Cuqlock-Knopp, V.G., Wilkins, C.A. and Torgerson, W.S. (1991). "Multiple Cue Probability Learning and the Design of Information Displays for Multiple Tasks." In D.L. Damos, Ed. Multiple-Task Performance, pp. 139-152, Taylor&Francis, London

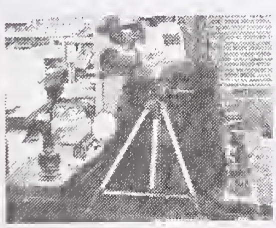
Fitts, P.M. (1964). "Perceptual-Motor Skill Learning." In A. W. Melton, Ed. Categories of Human Learning, pp. 243-285, Academic Press, New York, NY.

Rose, F; Attree, E.A.; Brooks, B.M. (2000) "Training in Virtual Environments: Transfer to Real World Tasks and Equivalence to Real Task Training." Ergonomics, v 43 no 4, Apr p. 494-511.

Sage, G. H. (1984). "Motor Learning and Control: A Neuropsychological Approach." Wm. Brown Publishers, Dubuque, Iowa.

Fig 1. Computer Integrated Backhoe at CARL

a) Motion Control



b) Visual Feedback (2D-3D)

Fig 3. Communication Hardware

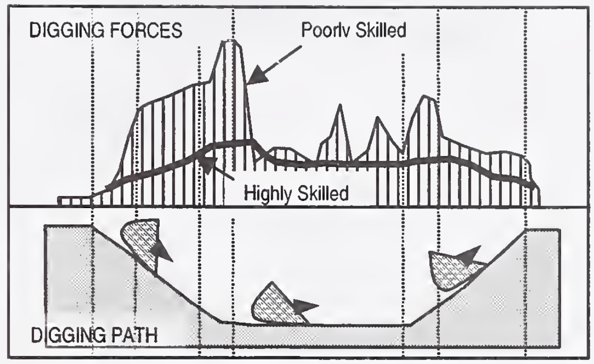


Fig 4. Comparison of Hypothetical Forces

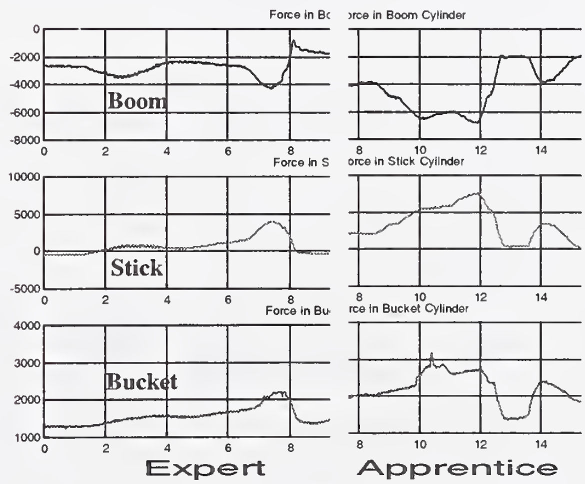


Fig 5. Measured Force Patterns

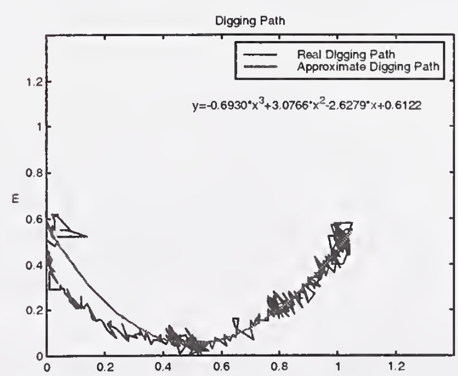


Fig 6. **Digging Path of "Apprentice" Operator**

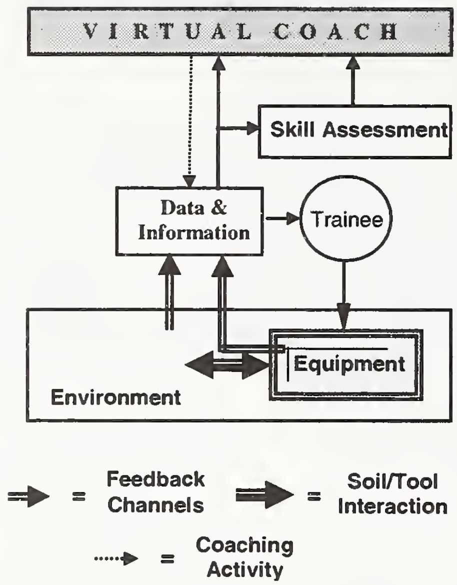


Fig 7. Layout of a Coached Training System

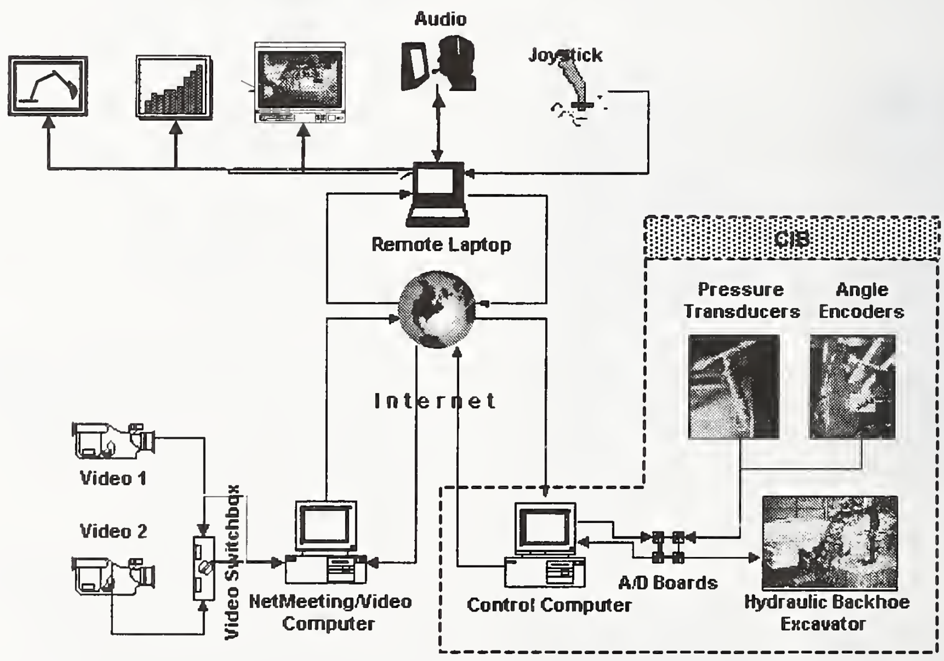


Fig 2. Schematic of Internet-Based-Operator-Training System

Terrain Aided Localization of Autonomous Vehicles†

Raj Madhavan
Intelligent Systems Division
National Institute of Standard and Technology
Gaithersburg, MD 20899-8230, U.S.A.
E-mail: **raj.madhavan@nist.gov**, **Fax: (301) 990-9688.**

Abstract— *This paper describes the development of a terrain-aided localization framework for autonomous land vehicles operating at high speeds in unstructured, expansive and harsh environments. The localization framework developed is sufficiently generic to be used on a variety of other autonomous land vehicles and is demonstrated by its implementation using field data collected from two different trials on two different vehicles. The results demonstrate the robustness of the proposed localization algorithms in producing reliable and accurate position estimates for autonomous vehicles operating in a variety of unstructured domains.*

Keywords— *Outdoor Localization, Map Building, Iterative Closest Point, Extended Kalman Filter, Entropy, Scale Space.*

I. Introduction

THE research addressed in this paper is concerned with the theoretical development and practical implementation of reliable and robust localization algorithms for autonomous land vehicles operating at high speeds in unstructured, expansive and harsh environments [1]. Localization is the ability of a vehicle to determine its position and orientation within an operating environment at any given time. The need for such a localization system is motivated by the requirement of developing autonomous vehicles in applications such as mining, agriculture, cargo handling and construction. The main drivers in these applications are safety, efficiency and productivity. The approach taken to the localization problem in this paper guarantees that the safety and reliability requirements imposed by such applications are achieved. The approach also aims to minimize the engineering or modification of the environment, such as adding artificial landmarks or other infrastructure, a key driver in the practical implementation of a localization algorithm [2].

In pursuit of these objectives, this paper develops a unified localization framework that uses measurements from both artificial and natural landmarks, combined with dead-reckoning sensors, to deliver reliable vehicle position estimates. The proposed localization framework is sufficiently generic to be used on a variety of other autonomous land vehicle systems. This is demonstrated by its implementation using field data collected from two different trials on two different vehicles. The first trial was carried out on a four-wheel drive vehicle in an underground mine tunnel. The second trial was conducted on a Load-Haul-Dump (LHD) truck in a test tunnel constructed to emulate an underground mine. The estimates of the proposed localization algorithms are compared to the *ground truth* provided by an artificial landmark-based localization algorithm that uses bearing measurements from a laser.

The paper is organized as follows: Section II develops an Iterative Closest Point - Extended Kalman Filter (ICP-EKF) algorithm - a map-based iconic algorithm that utilizes measurements from a scanning laser rangefinder to achieve localization. The ICP-EKF algorithm entails the development of a map-building algorithm. The development and implementation of an entropy-based metric to evaluate the information content of measurements and how this metric facilitates the augmentation of landmarks to the ICP-EKF algorithm guaranteeing reliable and robust localization is the subject of Section III. Section IV details the development and adaptation of a view-invariant Curvature Scale Space (CSS) landmark extraction algorithm. The algorithm is sufficiently robust to sensor noise and is capable of reliably detecting and extracting landmarks that are naturally present in the environment from laser rangefinder scans. The integration of the information metric, the CSS and the ICP-EKF algorithms to arrive at a minimal infrastructure localization framework is detailed in Section V. Finally, Section VI summarizes the key results.

II. Map-Based Iconic Localization

A bearing-only laser was mounted on the roof of the vehicles (for both the trials to be described in the following sections) so that it could detect strategically placed artificial landmarks (reflective stripes) in the trial environment. The exact position of these landmarks were made available from surveying using a digital theodolite. When the laser mounted on the vehicle moves through the environment, it detects the presence of these landmarks. Thus as the vehicle traverses through the environment, a sequence of bearing measurements to a number of fixed and known loca-

†The research reported in this paper was performed at the Australian Centre for Field Robotics, The University of Sydney, Sydney, Australia. The author would like to acknowledge the financial assistance provided towards his doctoral research by the Centre for Mining Technology and Equipment, Brisbane, Australia.

tions are made. Since the locations of these reflectors are known to the vehicle navigation system, the location of the vehicle can be computed from the bearing measurements made. Utilizing bearing measurements from a bearing-only laser in combination with dead-reckoning sensors (velocity and steering encoders and rate of change of orientation information from an inertial measurement unit), an EKF was employed to obtain ground truth.

The pose estimates obtained using the above artificial landmark algorithm were used to decouple the problems of map-building and map-aided pose estimation. Although it may not be appropriate to engineer the environment with artificial landmarks, it is possible to use a slow moving vehicle and a small number of landmarks to build a sufficiently accurate initial map. The reflective stripes required for the artificial landmark algorithm can be removed and the generated map can be used for vehicle localization. When the vehicle is stationary, it is easy to construct the vertices of a polyline that fit the data. When the vehicle starts moving, new range data obtained from the sensor needs to be integrated into the polyline map. Towards this, an incremental algorithm was developed for obtaining the optimal location of the vertices of the polyline in view of the new incoming data. Clearly as the vehicle moves, there will be observations that correspond to parts of the tunnel walls that were not seen before and therefore are not represented by the existing polyline. These observations were collected to expand the map by attaching line segments to either end of the polyline as the vehicle moves gradually along the tunnel. The proposed map-building algorithm is specific to this application and is not intended for high level tasks such as path planning where the required accuracy of the map would be higher. The polyline representation was chosen because of its simplicity. The curved surfaces of the environment can be represented using splines or a combination of line segments and arcs except for the fact that the additional complexity in such representations is not required for this application. As will be demonstrated, the maps obtained by the proposed map-building algorithm for both the 4WD vehicle and the LHD truck are adequate for localization.

Once the map is available, the next step towards achieving localization is map-registration. This stage is often referred to as the correspondence determination. Here, the correspondence problem involves registering the 2D laser range data to the 2D map. The Iterative Closest Point (ICP) algorithm [3] is employed to obtain the correspondence. As the iconic ICP algorithm works directly on sensed data, it does not require extraction and registering of features. The crux of the Iterative Closest Point (ICP) algorithm is to iteratively match points in one set to the closest points in another set, given that the transformation (the translation and/or rotation) between two sets is small. Here, the two point sets are the map and the laser range data, respectively. The shortcoming of the ICP algorithm is that it can only deal with cases when the first set is a subset of the second set. Zhang [3] developed a similar idea for establishing the correspondence which will be hereafter referred to as the ICP algorithm. The strength of this algorithm lies in the fact that it is capable of dealing with gross outliers in the data, occlusion and appearance and disappearance in which points in one set do not appear in the second set.

With the ICP algorithm there is no definitive way in which the uncertainty of the range data can be taken into account. Although Zhang discusses *partial* ways to accommodate the measurement uncertainty, ICP alone does not provide sufficiently reliable and accurate vehicle motion estimates. These shortcomings are overcome by combining the ICP with a post-correspondence EKF. The laser observations that do not correspond to any line segment of the map are discarded during the EKF update stage thus making it robust to errors in the map. Another attractive and appealing feature of this combined ICP-EKF algorithm is that observations from a variety of different sensors can be easily combined, since the changes are reflected only as additional observational states in the EKF [4], [5].

The vehicle employed in the first trial was a four-wheel drive (4WD) Troop Carrier shown in Figure 1. The estimated path of the 4WD vehicle (solid line) provided by the ICP-EKF algorithm along with the artificial landmark-based path (dotted line[1]) and the generated map using the proposed map-building algorithm is shown in Figure 2(a). Note that the map captures the geometry of the environment adequately. The curved surfaces are sufficiently modeled by shorter line segments in the polyline map. The vehicle travels a distance of 150 meters along the tunnel from right to left. The orientation estimated by the ICP-EKF (solid line) and that obtained by the artificial landmark algorithm (dotted line) are shown in Figure 2(b). The corresponding 2σ confidence bounds for the absolute error in x, y and ϕ are shown in Figure 3. It can be seen that the errors are bounded and thus the pose estimates are consistent. It is also clear that the estimated path is in close agreement with the artificial landmark-based path.

The second vehicle that was considered for the verification of the proposed algorithms is the LHD truck. LHDs are the work horses of the mining industry. The vehicle has a front and a rear body which can rotate

[1]As the estimates and the their corresponding ground truth are very close, extra effort is required on the part of the reader to distinguish between the two.

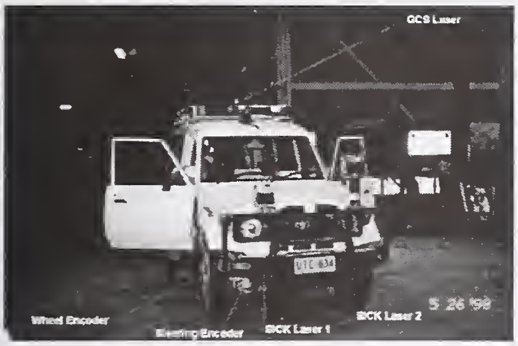
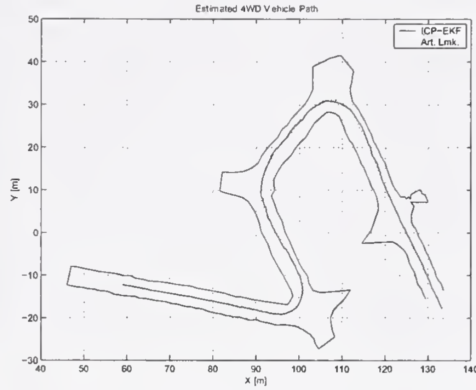
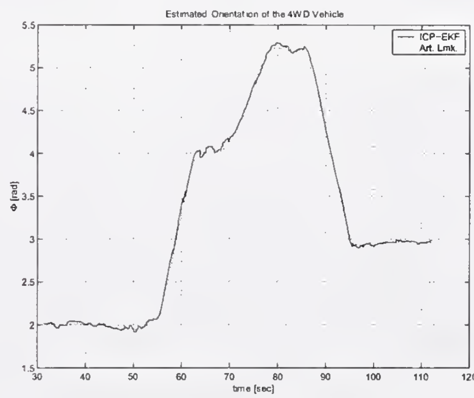


Fig. 1. The 4WD trial vehicle used in the underground mine trials. The location of the wheel and steering encoders, two time-of-flight range and bearing lasers and the bearing-only laser is shown.



(a)



(b)

Fig. 2. ICP-EKF estimated path (a) and orientation (b) of the 4WD vehicle (solid line) and the artificial landmark-based path and orientation (dotted line). In (a), the starting location of the vehicle is at $(133.3, -17.8)$ and the direction of travel is from right to left.

relative to each other and the front and rear wheel sets are fixed to remain parallel with the body of the vehicle. Steering is achieved by driving the articulation joint located mid-way between the front and rear axles. Both the front and rear wheel sets are driven at the same speed through a single transmission. The LHD and its sensor suite are shown in Figure 4.

Unique variations in data sets are necessary to establish an unambiguous correspondence with the map. Whenever the uniqueness can not be guaranteed, the ICP-EKF algorithm can fail to produce reliable association. The ICP algorithm provides a single transformation for each registration. Given the data sets, a set of transformations satisfies the registration. In the case of long tunnels or circular regions, this is too restrictive. For the tunnel, it is the set of transformations that align the walls independently of the position along the wall. Essentially what this means is that the position of the vehicle will be uncertain along the longitudinal direction of the tunnel. When using an EKF that combines range scan registration with dead-reckoning uncertainty, the positional covariance (or the uncertainty ellipse) will be large along the directions that can not be locked down by range scan registration alone. When the vehicle encounters a region that can be reliably recognized, the positional covariance can be reduced. When such regions are not encountered, some form of external aiding needs to be provided in regions where the ICP-EKF algorithm fails.

For the LHD, the ICP fails to produce correct correspondences in certain regions of the tunnel when there are degeneracies in rigid transformations during registration. The proposed method of overcoming these deficiencies is to incorporate landmarks that provide aiding information to guarantee reliable localization. A strategy to augment the ICP-EKF algorithm with a limited number of artificial landmarks is developed in Section V. These landmarks are then detected by the bearing-only laser to overcome limitations of the ICP-EKF algorithm.

Being an iterative algorithm, the initial pose estimate that is made available to the ICP is extremely important as the resultant correspondences depend on a good initial estimate. For the 4WD vehicle, the landmark augmentation procedure was not necessary in spite of long, straight sections of the tunnel. The reason for this is two fold. Firstly, the slippage the 4WD vehicle experiences due to the undulatory nature of the terrain is less than that of the LHD. In addition to the uneven nature of the terrain, the LHD slips much more than the 4WD vehicle due to the geometry of the tunnel and the kinematics of the truck

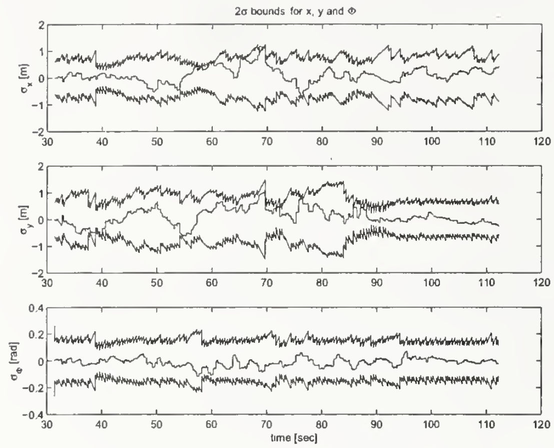


Fig. 3. The 2σ confidence bounds are computed using the covariance estimate for the error in x, y and ϕ compared to the actual error computed with the corresponding artificial landmark-based estimates for the 4WD trial.

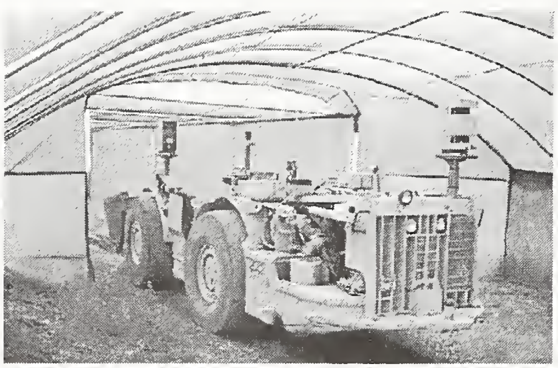


Fig. 4. The Load-Haul-Dump truck and its sensor suite. The LHD was fitted with two two-dimensional time-of-flight range and bearing lasers. The bucket is in front of the vehicle. A bearing-only laser scanner, as in the case of the 4WD trials, detects the presence of artificial landmarks. Encoders were used to measure the vehicle velocity and the articulation (steering) angle.

itself. Secondly, the 4WD vehicle model produces good prediction estimates enabling the ICP to provide accurate correspondences thus obviating the need for landmark augmentation. For the articulated LHD, dead-reckoning is not as good as that for the 4WD vehicle. Even explicitly taking into account wheel slippage with the inclusion of the two slip variables, α_v and β_v [1], the dead-reckoning estimates that are the initial estimates to the ICP are poor and accordingly the correspondences are no longer accurate. For these reasons, external landmark aiding becomes necessary for the LHD.

III. Landmark Selection and Augmentation

The proposed method for the selection of a particular landmark is based on localization information offered by a particular landmark from a given vehicle pose. This method implicitly takes into account the uncertainty in the vehicle pose estimate while computing the information content of the landmark. The concept of entropy is employed to facilitate landmark augmentation.

The entropy $h(\cdot)$ is a measure of the average uncertainty of a random variable and thus represents the *compactness* of the probability distribution. Subsequently, it is a measure of the *informativeness* of the distribution where *information* is defined as the negative of entropy. The entropy is minimum when information is maximum. It is conventional to seek minimal entropy when actually maximum information is sought. A mathematical expression for the entropy of a Gaussian distribution is to be developed. For an n-dimensional state vector $\mathbf{x}_k$ conditioned on a stacked observation vector denoted by $\mathbf{Z}_k \triangleq [z_1, z_2, \ldots, z_k]$ where $z_1, z_2, \ldots, z_k$ are individual sensor measurements, the *posterior entropy* can be derived to be [6]:

$$\begin{aligned} h_{k|k} &\triangleq h\left(p\left(\mathbf{x}_k \mid \mathbf{Z}_k\right)\right) \\ &= \mathrm{E}\left\{-\ell n\; p\left(\mathbf{x}_k \mid \mathbf{Z}_k\right)\right\} \\ &= 0.5\;\ell n\left[(2\pi e)^n \left|\mathbf{P}_{k|k}\right|\right] \end{aligned}$$

Thus for a Gaussian (normal) vector distribution all that is required to compute its entropy is its length, n and covariance, $\mathbf{P}$.

The *posterior* and *prior* information metrics can then be defined as:

$$\begin{aligned} im_{k|k} &\triangleq -h\left(p\left(\mathbf{x}_k \mid \mathbf{Z}_k\right)\right) \\ &= -0.5\;\ell n\left[(2\pi e)^n \left|\mathbf{P}_{k|k}\right|\right] \\ im_{k|k-1} &\triangleq -0.5\;\ell n\left[(2\pi e)^n \left|\mathbf{P}_{k|k-1}\right|\right] \end{aligned}$$

The resultant information contribution, ic, from measurements, is thus given by the relation:

$$ic_{k|k} \triangleq im_{k|k} - im_{k|k-1} \tag{1}$$

To overcome cases in which the ICP-EKF algorithm has insufficient information, the system is augmented with a limited number of artificial landmarks along certain sections of the tunnel. To realize this goal, the following three questions need to be answered: 1) When should a landmark or landmarks be added for external aiding? 2) How to select potential landmarks in order to overcome the divergence of the ICP-EKF estimates? Additionally, given a set of landmarks, how to select a landmark or landmarks that are optimal in the sense of increasing the robustness of localization? 3) Given a set of optimal landmark(s), how to incorporate these external aiding measurement(s) from the landmark(s) into the existing ICP-EKF framework?

The first two questions are answered in the following paragraphs and the landmark augmentation methodology is detailed in Section V.

• Landmarks are introduced when the estimated pose covariances exceed a predefined bound. The growth in covariance is a direct result of inaccurate correspondences provided by the ICP. Without additional landmarks, this would eventually lead to filter divergence.

The issue of computing a potential landmark set that can be employed to provide aiding measurements is addressed as follows:

• Potential landmarks can be computed from a given vehicle position since the artificial landmark locations are available as is the direction of vehicle travel. By discarding landmarks that are farther than a predetermined distance (proportional to the range of the bearing-only laser when artificial landmarks are to be augmented or to that of the range and bearing laser when natural landmarks are to be augmented), a landmark set can be arrived at and denoted by: $\mathbf{L}_{pl}$:($\mathrm{X}_{i_{pl}}$,$\mathrm{Y}_{i_{pl}}$), $i_{pl} = 1 : n_{pl}$, where n_{pl} is the number of landmarks in the set.

• From the current vehicle position, for all the landmarks in the set $\mathbf{L}_{pl}$, the predicted bearing measurement (range and bearing measurements in the case of natural landmarks) is computed and is checked to see if this measurement is acceptable.

• For all acceptable measurements, the individual information contributions are computed according to Equation (1). Finally, the landmark that provides the maximum information from the potential set of landmarks is selected for augmentation.

IV. Multiscale Natural Landmark Extraction

Feature or natural landmark-based methods for autonomous localization have become increasingly popular as they do not require infrastructure or other external information to be provided to the vehicle for operation. Feature-based localization methods require that natural landmarks can be robustly detected in sensor data, that the localization algorithm can reliably associate landmarks from one sensor observation to the next, and that the method can overcome problems of occlusion, ambiguity and noise inherent in measurement data. In indoor environments, features such as walls (line-segments), corners (diffuse points) or doorways (range discontinuities) are used as landmarks. In outdoor environments however, similar simple features are sparse and infrequently observed. In unstructured or natural outdoor environments a more general notion of a landmark or navigation feature is required. Given a (laser) range scan of an environment, one natural measure of feature or landmark significance is the local curvature in this range data. Rapidly changing range data indicates either a significant physical discontinuity or a prominent geoform. The use of local curvature as an interest measure is well known in the computer vision community. Further, and most important for navigation, curvature extrema are well known to be view-point invariant. Practically, this means that points of maximum curvature can be used as robust point landmarks in localization.

Scale space filtering is a qualitative signal description method that deals with the problem of *scale* by treating the size of a *smoothing kernel* as a continuous parameter [7]. The main idea of multiscale representation is to successively suppress fine scale information and thus to progressively remove the high frequency information which results in the signal becoming gradually smoother. The common way of constructing a scale space is by convolution with a Gaussian kernel. The essential requirement here is that a signal at a coarser (higher) scale level should contain less structure than at a finer (lower) level of scale. In scale space methods, a qualitative description of the signal is obtained by studying the behavior of the extrema over the continuum of scales. Also the extrema that survive over larger smoothing extents are considered to be more significant than others. Attenuation of the noise in the signal is realized by convolving the signal with the Gaussian kernel.

A Curvature Scale Space (CSS) algorithm was developed to identify, extract and localize landmarks characterized by points of maximum curvature at successive geometric scales. The CSS algorithm can be used to extract curvature extrema from laser scan data corresponding to landmarks at different scales that are: 1) invariant to rotation and translation of the shape (signal) under consideration 2) robust to noise and 3) reliably detected and localized. The basic principle of the CSS algorithm to identify dominant points is: "Convolve a signal with a Gaussian kernel and impart smoothing at different levels of scale (the scale being proportional to the width of the kernel). From the resulting convolved signal, identify the dominant points (curvature extrema)." For various degrees of smoothing of the curve (segmented range scan), it is desired to find the curvature extrema. A parametrization of the curve is necessary to compute the curvature at varying levels of detail (for segmentation and parametrization procedures see [8]). A parametrization is possible by considering a path length variable along the curve and expressing the curve in terms of two functions $x(s)$ and $y(s)$ such that $C = \{x(s), y(s)\}$ with s being a linear function of the path length ranging over the closed interval $[0, 1]$. The curvature, κ, is given by:

$$\kappa(s,\sigma) = \dot{X}(s,\sigma)\ddot{Y}(s,\sigma) - \dot{Y}(s,\sigma)\ddot{X}(s,\sigma)$$

where $\left\{\dot{X}(s,\sigma), \dot{Y}(s,\sigma)\right\}$ and $\left\{\ddot{X}(s,\sigma), \ddot{Y}(s,\sigma)\right\}$ are obtained, respectively, by convolving $x(s)$ and $y(s)$ with the first and second derivatives of the Gaussian kernel. The Gaussian kernel is given by: $g(s,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-s^2/2\sigma^2}$ where σ is the width of the kernel. By employing the CSS algorithm, dominant points are extracted based on the persistence within a scan and over consecutive scans. Features that exist only at finer scales and disappear at higher scales do not correspond to stable features and thus do not qualify as reliable candidates for subsequent detection and tracking during the vehicle localization stages. Persisting dominant points enable the construction of a natural landmark map based on which natural landmark-based localization is achieved.

V. Landmark Augmented Minimal Infrastructure Localization

This section describes the localization of the LHD using both artificial and natural landmarks. Natural landmarks characterized by dominant points of curvature were extracted from laser rangefinder scans and a natural landmark map was built using the curvature scale space algorithm as detailed in Section IV. Such a natural landmark map can then be used to augment the ICP-EKF framework thereby reducing the number of artificial landmarks that are required for reliable and robust localization.

A. Natural Landmark Observation Model

The laser rangefinder provides both range and bearing to a landmark and accordingly an equation relating both of them to the vehicle (laser) location is required. The predicted range and bearing, respectively, for each natural landmark j at discrete time-instant k is given by the non-linear model:

$$\begin{aligned} R^j_{n\ell_k} &= \sqrt{\left[x^j_{n\ell} - x_{L_k}\right]^2 + \left[y^j_{n\ell} - y_{L_k}\right]^2} \\ \theta^j_{n\ell_k} &= \tan^{-1}\left[\frac{y^j_{n\ell} - y_{L_k}}{x^j_{n\ell} - x_{L_k}}\right] - \phi_{v_k} \end{aligned}$$

where $(x^j_{n\ell}, y^j_{n\ell})$ is the cartesian location of landmark j and (x_L, y_L) is the location of the laser rangefinder on the vehicle. The observation model for a natural landmark is thus given by:

$$Z^j_{n\ell_k} = \begin{bmatrix} R^j_{n\ell_k} \\ \theta^j_{n\ell_k} \end{bmatrix} + \begin{bmatrix} v^R_{n\ell_k} \\ v^\theta_{n\ell_k} \end{bmatrix} \tag{2}$$

where $v^R_{n\ell_k}$ and $v^\theta_{n\ell_k}$ refer to the uncertainty present in the range and bearing measurements and are modeled as zero-mean uncorrelated Gaussian sequences with constant variances, $\sigma^2_{R_{n\ell}}$ and $\sigma^2_{\theta_{n\ell}}$, respectively.

B. Entropy-based Artificial Landmark Selection and Augmentation

During the vehicle localization stage, for every range and bearing scan, natural landmarks are extracted and are associated with those in the map. Details of natural landmark matching and discriminance procedures are detailed in Section V-C. When the information provided by these natural landmarks is not enough to curtail the growth of the pose covariances, the entropy-based landmark selection and augmentation method is adopted to select the required artificial landmarks. Whenever it is determined that the ICP-EKF framework requires additional external aiding, all the landmarks that are visible from the current vehicle position are examined and the landmark that contributes the maximum information is then selected.

C. Natural Landmark Matching and Discriminance

Whenever a range and bearing scan becomes available, there is the need to decide whether the scan should be used for a) Natural landmark-based localization or b) ICP-EKF based localization. The same scan should not be used for both as this amounts to using the same information twice. To avoid such reuse of information, the entropy-based information content measure is utilized to determine to decide upon (a) or (b). The pertinent localization procedure is arrived at by looking at the information contribution provided by both (a) and (b) and selecting the one that provides maximum information towards localization.

When natural landmarks are used for localization, it is essential that the extracted natural landmarks from a scan are reliably matched to those in the map. Whenever a range and bearing laser scan is received, the dominant point landmarks are extracted in the same fashion as described in Section IV. Locations of all the extracted dominant points above a certain scale level are then compared with the natural landmark map. Two landmarks that differ only slightly in their position should be identified as distinct within the limits imposed by the sensor noise. Such discriminance is achieved by checking if the computed innovation sequences pertaining to both the range and bearing of a matched natural landmark fall within the normalized innovation gate by using the natural landmark observation model given in Equation (2).

If the observations fall within the prescribed limits, then the landmark is accepted as a match. Subsequently, this observation is used to update the vehicle states using the range and bearing innovation sequences that are obtained as a direct consequence of the matched landmark measurement. The resultant information contribution from a matched natural landmark, $ic_{n\ell}$, and that after the ICP-EKF updates for the same scan, $ic_{icp-ekf}$, are computed by using Equation (1). Based on $ic_{n\ell}$ and $ic_{icp-ekf}$, the procedure that provides the maximum information is chosen. The combined algorithm thus bounds the natural degradation of the ICP registering results when there are no distinguishable landmarks along the longitudinal direction and ensures that the lateral matching (crucial for vehicle guidance) is not deteriorated.

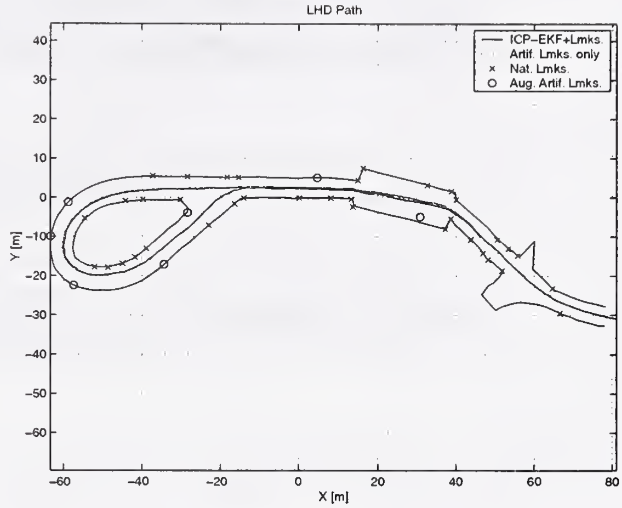
D. Results and Discussion

The results when both natural and artificial landmarks are employed towards localization of the LHD are discussed now. Figure 5(a) and (b) show the estimated path and orientation of the LHD by the ICP-EKF natural and artificial landmark-augmented algorithm (solid line). It should be emphasized here that the natural landmarks need to be extracted only once. For localization in a tunnel over several trials, the natural landmark extraction can be done on an exploratory trial run and such landmarks can then be utilized. From Figure 5(a), it can be seen that the landmark-augmented ICP-EKF algorithm provides estimates that are similar to the ground truth. Note that natural landmarks have taken the place of artificial landmarks along the straight sections of the tunnel. Unfortunately, in the circular loop section of the tunnel, there were not many persistent extractable natural landmarks and this necessitated the inclusion of artificial landmarks. Nevertheless, the important point to note here is that the number of required artificial landmarks have been successfully reduced (by a factor of four [1]). In environments where the extraction of persistent dominant landmarks is possible, the artificial landmarks can be totally eliminated[2]. Figure 5 (b) shows the orientation of the two trials (solid line) along with those provided by the artificial landmark algorithm (dotted line).

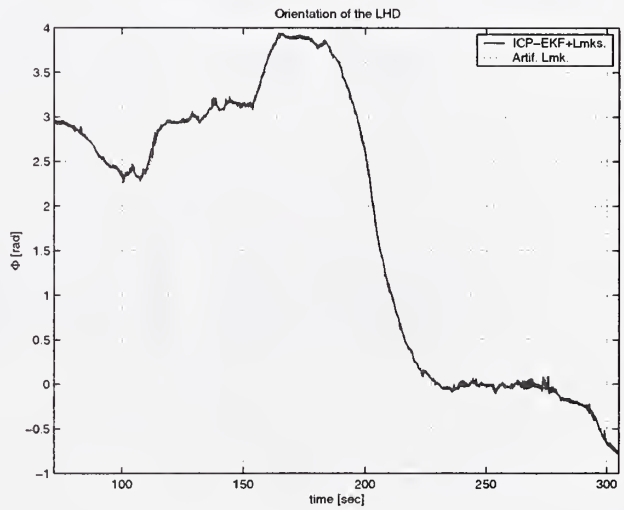
Figure 6 shows the absolute pose errors computed with the aid of the ground truth and the corresponding 2σ (95%) confidence bounds. It is evident that the errors remain white and are well within with the prescribed bounds representative of consistent estimates. Absolute error estimates of half a meter were achieved in the position estimates with the corresponding error in the orientation estimate being less that 5 degrees.

Figure 7 depicts the validated range and bearing innovation sequences along with the 1σ and 2σ bounds. The discontinuity in the innovation sequences corresponds to periods where there were no updates performed as there were no natural landmarks available during such periods. For example in Figure 7, the discontinuity from 177 – 200 seconds in both the range and bearing innovation sequences corresponds to the loop section of the tunnel where the external aiding is predominantly provided by the artificial landmarks. It can be seen that both the range and bearing innovation sequences also remain white and are well within the defined bounds.

[2]In fact, this claim was substantiated by a third trial run on a Utility vehicle in an outdoor area populated by people and moving cars in the University of Sydney campus [8].



(a) Estimated path of the LHD



(b) Estimated orientation of the LHD

Fig. 5. Natural landmark-augmented ICP-EKF estimated path (a) and orientation (b) of the LHD (solid line) and the artificial landmark-based path (dotted line). In (a), the traversed path runs right to left. The direction of travel is in the anti-clockwise direction around the loop at the left end of the figure. The circles represent the augmented artificial landmark locations and the crosses represent the natural landmark locations. The starting location is at (81.1, −30.9).

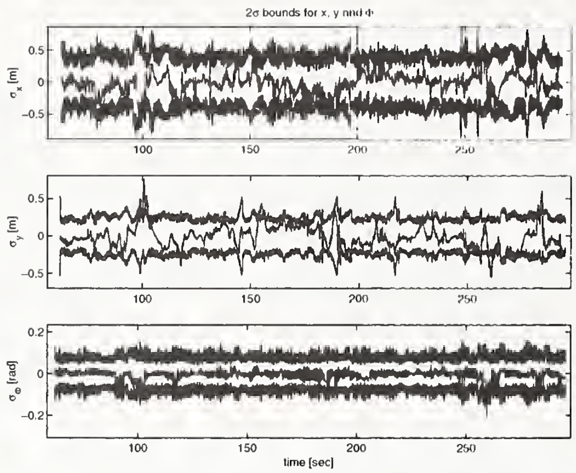


Fig. 6. Natural landmark-augmented ICP-EKF pose errors and the 2σ confidence bounds computed using the covariance estimate for the error in x, y and ϕ compared to the actual error computed with the corresponding artificial landmark-based estimates.

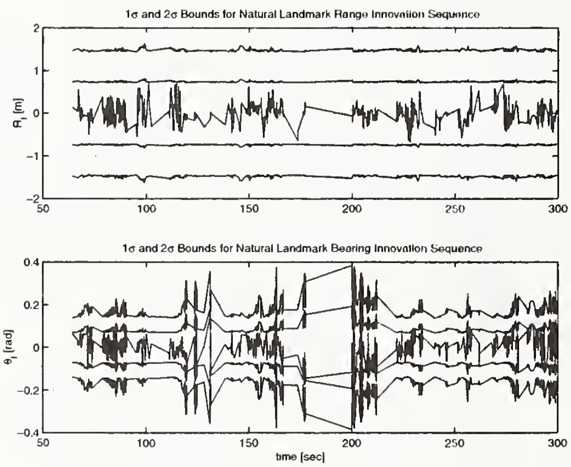


Fig. 7. Range and bearing innovation sequences for the natural landmarks matched during the LHD trial. The 1σ and 2σ confidence bounds are also shown.

VI. Conclusions

This paper discussed the development of a unified terrain-aided framework for achieving minimal infrastructure localization of high speed vehicles operating in unstructured and harsh environments.

The first step in the achievement of this goal was the development of a map-based ICP-EKF localization algorithm utilizing measurements from a scanning laser rangefinder in combination with dead-reckoning sensors. The next step was the identification of shortcomings of the ICP-EKF algorithm and the development of an entropy-based landmark augmentation metric to overcome the deficiencies. Using this metric, evaluation of information content of measurements was possible and thus facilitating the acceptance or rejection of a particular landmark measurement. The metric was shown to be an optimal way of efficiently utilizing measurements by implicitly incorporating the landmarks' utility towards reducing localization error. A view invariant multiscale natural landmark extraction algorithm was developed to extract point landmarks characterized by dominant curvature points from laser scans thus facilitating reliable detection and extraction of persistent natural landmarks. Attenuation of noise inherently present in laser scans was achieved by convolving the scans with a Gaussian kernel. The algorithm was also shown to possesses the attractive property of being invariant to rotation and translation effects that the laser scans under consideration might experience due to vehicle travel over harsh terrains.

Finally, the paper detailed the integration of the information metric, the CSS and the ICP-EKF algorithms to arrive at a unified localization framework. The developed localization framework has the ability to use measurements from both artificial and natural landmarks as and when they become available and was shown to be sufficiently generic to be used on a variety of autonomous land vehicles, by its application to a 4WD vehicle and an LHD truck. The results demonstrated the reliability and robustness of the proposed framework.

References

[1] R. Madhavan, *Terrain Aided Localisation of Autonomous Vehicles in Unstructured Environments*, Ph.D. thesis, School of Aeronautical, Mechanical and Mechatronic Engineering, The University of Sydney, Jan. 2001.

[2] R. Madhavan, G. Dissanayake, H. Durrant-Whyte, J. Roberts, P. Corke, and J. Cunningham, "Issues in Autonomous Navigation of Underground Vehicles," *Special Edition of the Journal of Mineral Resources Engineering*, vol. 8, pp. 313–324, 1999.

[3] Z. Zhang, "Iterative Point Matching for Registration of Free-Form Curves and Surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 119–152, 1994.

[4] R. Madhavan, G. Dissanayake, and H. Durrant-Whyte, "Autonomous Underground Navigation of an LHD using a Combined ICP-EKF Approach," in *Proceedings of the IEEE International Conference on Robotics and Automation*, May 1998, pp. 3701–3708.

[5] R. Madhavan, G. Dissanayake, and H. Durrant-Whyte, "Map-Building and Map-Based Localization in an Underground-mine by Statistical Pattern Matching," in *Proceedings of the IEEE International Conference on Pattern Recognition*, Aug. 1998, pp. 1744–1746.

[6] J. Manyika and H. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*, Ellis Horwood Limited, 1994.

[7] A. Witkin, "Scale-space Filtering," in *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Aug. 1983, pp. 1019–1022.

[8] R. Madhavan, H. Durrant-Whyte, and G. Dissanayake, "Natural Landmark-based Autonomous Navigation using Curvature Scale Space," in *Proceedings of the IEEE International Conference on Robotics and Automation*, May 2002.

MONITORING OF THE BORING TRAJECTORY IN UNDERGROUND CHANNEL

Thomas Bock, Alexej Bulgakow*, Dimitry Krapivin, Sergej Aleksyuk, Sergej Pritchin**

**Technical University Munich, a.bulgakow@gmx.de*
*** South Russia State Technical University, aprim@novoch.ru*

Abstract: This paper deals with high precision directional boring of underground conduits, including programmable curved trajectories for enclosed digging driving communications and other purposes.

Keywords: underground channel, dynamic monitoring, driving communisations.

1. INTRODUCTION: As the working tool for boring in soft anisotropic soils a cone is modeled, which is imparted a revolving motion whose angular velocity is ω_{pr}, its impact frequency f. The impact makes the revolving cone penetrate into the soil along the screw trajectory. The cone thus revolves around its symmetry axis OO_1 with angular velocity ω_K

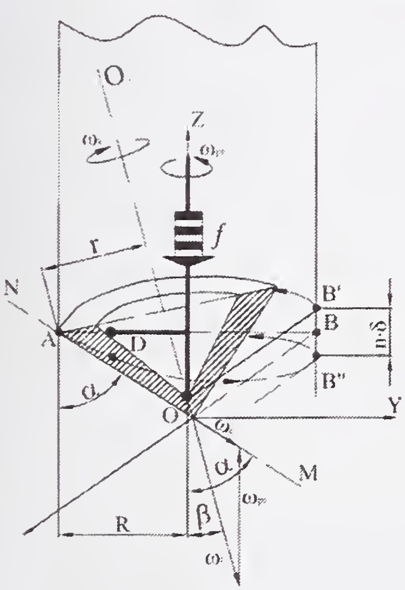


Fig. 1

(Fig. 1) in such a way that the above axis OO_1 itself oscillates eccentrically around O_Z at an angular velocity ω_0 at an angle β.

Acted on by various factors the conditions of oscillating motion of the OO_1 axis can change

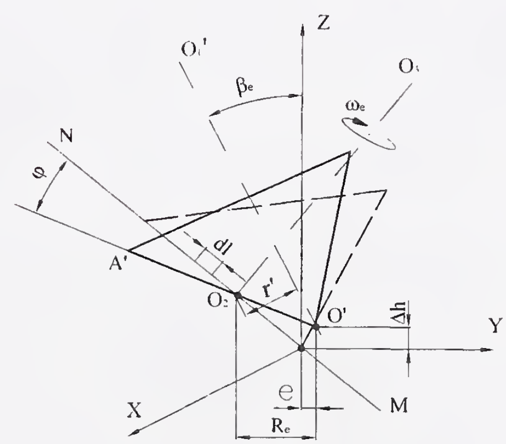


Fig. 2

in such a way that the apex of the cone is displaced from the axis of the O_Z, relative to the tunnel centerline and takes part in oscillating motion as well, its parameters being e and Δh to O_Z axis (Fig. 2). Here, the cone slips and performs friction work while its constituent OA is displaced from NM line during its turn at some angle φ to O_2 and O_3 at an angular velocity ω_e.

The above parameters of this motion are found following the calculation of the location of point O_2, based on the principle of the least action. With uniform distribution of the normal force longitudinal to the generating line of OA, the elementary work of the friction force F on the element $d\ell$ will be

$$dA = F \cdot V \cdot d\ell,$$

where $V = \omega_\varepsilon \cdot S$ = the linear rate of $d\ell$ element, and S = the distance of $d\ell$ element up to NM.

The total work of the friction forces is

$$A = \frac{F}{2} \cdot \omega_e \cdot \frac{\cos\beta \cdot \sin\varphi}{\sin(\beta-\alpha)} \cdot \left[(L-\ell)^2 + \ell^2\right].$$

The minimum of the function A will be at $\frac{dA}{d\ell}=0$ or with $\ell = \frac{L}{2}$, where L=OA, i. e. when the cone constituent slips along the tunnel boring face relative to point O_2 located in the middle of cone assembly.

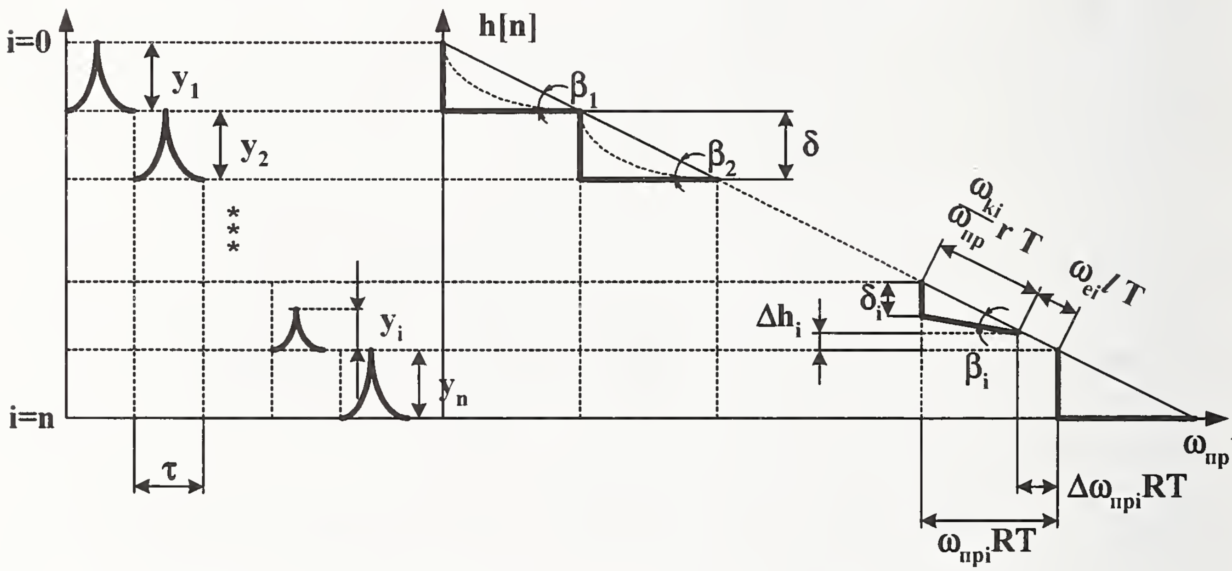


Fig. 3

The slip appearance will make point A shift from the screw line B'AB". The coordinates of point A can be derived as

$$x = (R - e) \cdot \cos\omega_{pr}t;$$
$$x = (R - e) \cdot \sin\omega_{pr}t; \qquad z = n \cdot \delta,$$

where n = number of impacts, δ = penetration in a direction of operation of impacts; the R-e value is found from

$$R - e = \frac{\omega_k}{\omega_{pr}} \cdot r,$$

where r is the greatest cone diameter.

Thus, dynamic monitoring of the trajectory of a working cone motion is carried out by means of defining angular velocities ω_e, ω_{pr} and analysis of the amplitude readings of impacts. These indications can be used to control the directed cone motion. The shift value, e, at every turn can change with varying ω_{pr} and redistributing the number of impacts around the face perimeter.

When boring the tunnel in a uniform medium e is a constant. In a particular case, if e=0, the motion parameter relation is observed

$$h[n] = \frac{\omega_{\kappa} \cdot r}{\omega_{pr}} \cdot \sin\beta, \qquad (1)$$

Where h[n] is the boring length in the preset direction within n impacts, where h=nδ, and β is the angle of screw line inclination. Figure 3 shows the plane projection of a screw line for a turn.

The impact of duration τ=T and amplitude y causes line B`AB" to be formed, which consists of curved portions of separate tunneling for every impact, thus resulting in screw line inclination at angle β.

If the medium is non-uniform, the parameter relationship in equation (1) is violated. Thus, in Figure 3 at the i-th stage the impact amplitude $y_i \ll y_n$, the impact duration $\tau_i < T$, but Δh_i decreases $\delta_i < \delta$, the angle $\beta_i < \beta$ changes and there appears eccentricity e_i. All the above can result in trajectory deviation due to variation of dynamic properties of the tunneling process.

To correct this deviation operatively one must have a self-adjusting system of automatic controlling of the direction, which is effective if the anisotropy of the medium acts weakly. The system shown figure 4 has as a source of data gauge 1 of the amplitude of impact, gauge 2 of angular velocity of the turning cone ω_{κ}, gauge 3 of the angle of drive rotation velocity, a discrete correlator, a controlled filter, and an adaptive regulator. To find the dynamic properties of the system consisting of a drive feeding set, and a penetrating cone with a striker, a pulse transient function as a sequences of coordinates is synthesized in a controlled filter as:

$$W_0 = W(0)\,,\ W_1 = W(\tau)\,,\ W_2 = W(2\cdot\tau)\,,$$
$$W_n = W(n\cdot\tau).$$

These values enter the RAM on OS_1.....OS_n.

Calibration values $\omega^*_0 - \omega^*_n$ are recorded on OS^*_1 to OS^*_n to store the programmable trajectory into ROM.

On the delay line a periodic signal is put from the input unit, which is proportional to the centered correlation input – output function. Thus, the following expression is obtained on the adder output:

$$K^{*0}_{xy}(\tau) = T_m \cdot \sum_{i=0}^{n} W(i\cdot T_m)\cdot K^0_y\cdot(\tau - iT_m)\,.$$

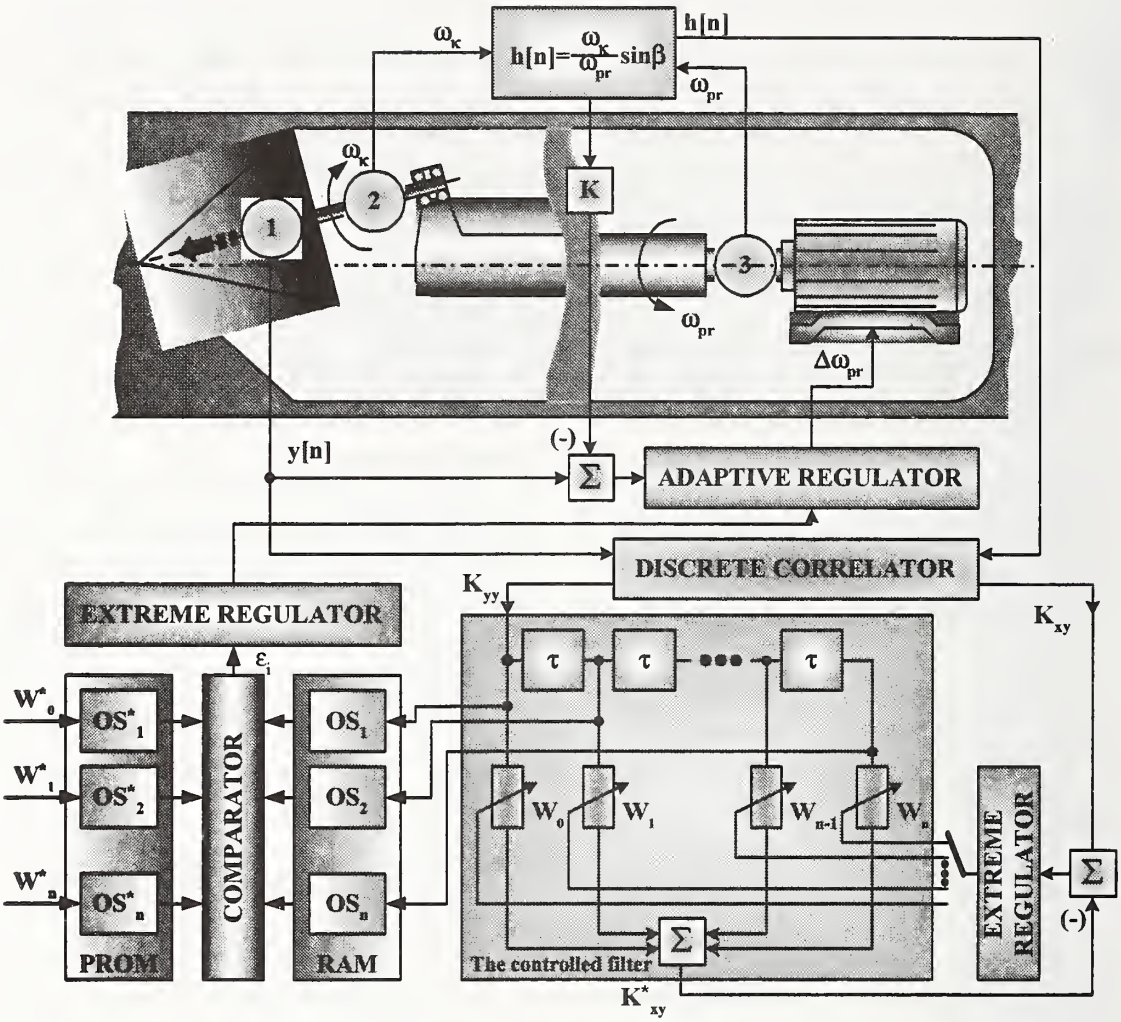
The difference is put to the comparator:

$$\varepsilon = K^{0*}_{xy}(\tau) - K^0_{xy}(\tau),$$

By adjusting the coefficients this difference can be minimized. Extreme regulators are used to make the process of impact function estimate automatic.

2.0 CONCLUSIONS:

Thus, the procedure of trajectory monitoring envisages a complete analysis of dynamic system properties taking into account stochastic character of effect. This control method uses intensifying impact effects for probing anisotropic properties at the face and controlling the process of direction correction.



h[n]
ω_K
$h[n]=\frac{\omega_K}{\omega_{pr}}\sin\beta$
ω_{pr}
ω_K
2
1
K
3
ω_{pr}
$\Delta\omega_{pr}$
(-)
y[n]
ADAPTIVE REGULATOR
DISCRETE CORRELATOR
K_{yy}
K_{xy}
EXTREME REGULATOR
ε_i
W^*_0
W^*_1
W^*_n
OS^*_1
OS^*_2
OS^*_n
COMPARATOR
OS_1
OS_2
OS_n
PROM
RAM
W_0
W_1
W_{n-1}
W_n
The controlled filter
K^*_{xy}
EXTREME REGULATOR
(-)

Fig. 4

SESSION 9

AUTOMATED INSPECTION AND MAINTENANCE MANAGEMENT SYSTEMS

A MULTI-SENSORY APPROACH TO 3-D MAPPING OF UNDERGROUND UTILITIES

Leonhard E. Bernold,
Associate Professor
Lashminarayan Venkatesan, *GRA*
Sushil Suvarna, *GRA*
Construction Automation Robotics Laboratory (CARL), North Carolina State University.

Abstract: The demand for new buried utilities, such as gas, power and fiber-optic communication lines is growing with new construction, re-construction, and the growth of the communication infrastructure worldwide. Because the machinery for placing the new utilities underground, such as backhoe excavators, trenchers, augers, drills, and plows, don't "feel" when they are getting close to already buried object, utilities are easily damaged. Despite great efforts in locating existing utilities before a contractor is allowed to dig, accidents occur. This paper will present a novel technology for detecting and locating buried utilities that attaches to the digging equipment and utilizes both EMI (Electromagnetic Induction) and a GPR (Ground Penetrating Radar). The newest effort involves the development and performance analysis of algorithms to detect and extract the features and characteristics of these utilities, such as their orientation, diameters etc. One prime focus is to minimize the percentage of false alarms. For that purpose, the two sensor systems are fused to create a multi-sensory approach to 3-D mapping of all the utilities without a priori knowledge of their location

Keywords: Buried utility detection, Electromagnetic Induction, Ground Penetrating Radar, Sensor fusion.

1. INTRODUCTION

In 1993, the Construction Automation and Robotics Laboratory (CARL) at North Carolina State University (NCSU) started an initiative to address the national problem of detecting and locating underground buried utilities. The core idea was to provide the equipment operator with his own system integrated with the equipment that alerts him of the danger rather then to depend on the color marks made by the locator sent by the One-Call center. For that purpose, a sensing platform, operating like a subsurface "X-ray" was attached directly to the machinery hence providing the operator with an opportunity to "see" and be warned when the machine tool gets close to an existing utility. The original system was based on the Electro Magnetic Induction (EMI) technology that was integrated with PC-based software to process and analyze the signal coming from an antenna. Using a traditional coil, the antenna generates its own magnetic field and senses the existence of ferrous and non-ferrous material. The analogue output of the controller is then digitized and plotted on the computer screen. By taking advantage of this capability, the technology was used to retrofit backhoe excavators, trenchers, and augers. The success of the initially crude system led to the development of an improved version to be used during the excavation of unexploded ordinances (UXOs). In 1997, the Buried Utility Detection System (BUDS) consortium was founded that continued the work on three fronts: a) processing of sensory data, b) mechanical system, and c) human-machine interface. A simple user interface was created by creating a control box for the operator, consisting of two buttons for operating the articulated sensory platform and a red-yellow-green light as feedback. Details of the system and current developments, experiments and preliminary results will be discussed later.

Utilities, such as gas lines and optical communication cables, are mainly non-metallic making the EMI technology useless since these utilities are "invisible" to the metal detector. Other non-metallic objects underground include sewer lines made of concrete, clay and plastic. One sensor that has been successfully used even in archeology is the Ground Penetrating Radar GPR). The GPR transmits RF signals and detects the signals reflected by changes in the ground. When translated on the surface it provides a cross-sectional image of the material below the ground surface. Details of how this image is decoded, problems faced and implementation details are discussed later.

2. THE NEED FOR DAMAGE PREVENTION

The congressional Transportation Equity Act for the 21st Century, TEA 21, Title VII, Subtitle C, SEC. 87301, states that: "...unintentional damage to underground facilities during excavation is a significant cause of disruptions in telecommunications, water supply, electric power, and other vital public services, such as hospital and air traffic control operations, and is a leading cause of natural gas and hazardous liquid pipeline accidents."

Underground Focus Magazine (1999) is a source that publishes an Accident File in every issue. For example it listed that from December 8th until December 11th 1998, seven major accidents occurred. On the 9th a fiber optic cable was cut by an excavation contractor that supported the 911 service for five counties in Jacksonville, Texas. The most tragic accident, however, occurred on Dec. 11 when "a crew using an "anchor cranker" to install a guy wire anchor for a telecommunications pole augured into a gas main." four people were killed and fourteen injured when the gas exploded in St. Cloud, MN.

There are many different parties, actively and passively, involved in the excavation and trenching process. Active participants include 1) owners of a new facility, 2) designers, 3) planners, 4) contractors, 5) utilities, 6) locators, 7) construction workers, and 8) equipment operators. In most U.S states, a contractor is required by law to call a "One-Call Center" 48 or more hours before he digs.

Despite the successful implementation of the One-Call systems in most of the U.S states, the accidents caused by damaging underground utilities resulting in wide variety of impacts reaching from a clogged residential sewer line to a gas explosion causing death and destruction. The list of impacted parties that incur cost comprises not only the contractor, utility and property owners, people in the vicinity of the accidents, but also the customers of a disrupted utility. Some of these groups include: a) private homes, b) governmental agencies, c) service companies, d) schools, e) hospitals, f) industrial firms, g) transportation systems like airports, taxi services, freight trains and trucking, h) retailers, and i) the utilities themselves. Overall, the direct and indirect costs of such accidents are staggering making the use of more sophisticated prevention approaches also economically prudent.

3. LOCATING BURIED METALLIC UTILITIES USING EMI

3.1. CONCEPT AND BACKGROUND

Anything metallic present in the ground can be induced to create a magnetic field, which can be detected by an antenna. The magnetic field is caused by a signal emitted by the transmitter coil. A receiver coil "listens" to this reflected signal and gets a measure of the metal around it. This can be done by either Continuous wave EMI or Pulse EMI.

EMI sensors have been used since the 1950's for quality control on manufacturing production lines to safeguard against contamination. In more recent times, they have been used as a tool for mining, non-destructive testing, security, archaeology geology and other related fields. Metal detectors using electro-magnetic induction, especially pulse induction are not new in the field of buried utility detection. With two antennae, or one moving antenna in many positions, it is also possible to determine the depth of the buried pipe. Some researchers (Das et al., 1990) performed the analysis of the EMI detector for real-time location of buried objects. Several response characteristics of the utility, such as object depth, orientation, aspect ratio, and material properties were studied and, due to limitations of direct metal detection, the need for sophisticated processing was observed. This technology has also been used to discover unexploded ordinance (Lorenc and Bernold, 1997).

3.2. A BURIED UTILITY DETECTION SYSTEM

Figure 1 presents the basic idea of the equipment-mounted Buried Utility Detection System (BUDS) developed at NC State University. It is a prototype aimed at being part of a real-time system, integrating sensor fusion techniques. Presently, efforts are being made to form an extensive database of utility contours obtained from BUDS. After forming such a database, the next step is to go in for field-testing of unknown soils and utilities buried underneath. With a knowledge base and classification based learning schemes, we would be in a position to estimate with a degree of probability, the depth, orientation and material properties of the object. This estimation or the data obtained in the cruder stage of the sensing process would be fed into the sensor fusion module.

3.3. EXPERIMENTAL TEST BED

The experimental setup developed at CARL, called BUDS, consists of a moving cart under a stationary antenna. The cart and the antenna can be positioned at various angles to represent real-life site configurations. It is completely software controlled and is being used for collecting sensor contours for various geometric configurations of pipes, changing the material properties of the pipes and the antenna characteristics itself.

3.4. INITIAL RESULTS

Initial results (Fig 2) from the experiments show predictable agreement with previous work in the field. It has been found that as the antenna scans the utility below and the computer continuously plots the readings, a peak in the graph shows strong metallic content very near the antenna. By changing the horizontal angle and vertical tilt we get slightly different curves and by triangulation we can estimate depth of the pipe. We also noticed the change in the curves for deriving material characteristics, for example, studying the difference in the contours for a solid pipe and hollow pipe of the same dimensions and material.

3.5 APPLICATIONS

Apart from just carrying out experiments for detecting utilities, the CARL team has also worked on a BUDS application. It consists of the antenna mounted on a backhoe giving feedback to the operator about the existence of buried utilities before excavating (Fig 3).

4. SUBSURFACE UTILITY MAPPING USING GPR

4.1. WHAT IS A GPR?

The GPR is a remote sensing short-range system, which measures short pulse electromagnetic (EM) reflections due to variations of the electrical properties of the investigated medium. The electromagnetic wave, which is radiated from a transmitting antenna, travels through the material at a velocity that is related to the electrical properties of the material. As the wave propagates, if it hits an object or a boundary with different electrical properties, then part of the wave energy is reflected or scattered back to the source. The wave, that is reflected back, is captured by an antenna and an image is created that is reflective of the materials and boundaries present beneath the surface. The main drawback of a GPR is the inability to detect the exact material of the buried object.

4.2. GPR USED IN PIPE AND MINE DETECTION

New and general methods for landmine detection using GPR images have been evaluated. Simple and effective observation vector representations have been constructed to model the time varying signatures produced by the interaction of the GPR and the landmines. Gader (et al., 2001) used Hidden Markov Models (HMM) to recognize patterns and to predict the presence of landmines. Landmines appear in time domain GPR as shapes similar to hyperbolas, although corrupted by noise and other factors. A signature library was created using a combination of ground truth and GPR response for that truth-value. HMMs were used to generate probabilities for the unknown images by comparing them with the signature library present. The GPR has also been used for pipe detection. The same concept of creating signature libraries for different pipes at different depths is used and searching algorithms are used to predict the occurrence of pipes from unknown images. The image consists of distinct patterns, e.g. a hyperbola, which are studied to obtain a result.

4.3. LABORATORY SETUP

Most GPR image processing algorithms are based on a signature database that maps the different possible objects with their orientations and the images created by these objects under different soil depths and conditions. The optimal algorithm processes the image of the unknown object and compares it with those present in the database and generates an approximate estimation about the nature of the object. To study the responses of different pipe materials and other objects that could be present underground an experimental workspace (Fig 4) has been setup.

Most sample tests performed in the detection of mines or pipes, involved moving the GPR in one line, either forward or backwards. However, our setup aims at observing images in a single plane initially, and evolving into one that considers the perpendicular movement of the GPR. This would result in an overall zigzag movement that would generate a 3-D GPR image.

4.4 CURRENT WORK / RESULTS

The setup shown in Figure 4 is being used to study patterns generated by pipes of different materials, kept at various depths and also patterns generated by objects such as rocks and wood that might be present. The responses from various materials have been collected and the images are being studied to locate patterns unique to each object. The next stages include studying patterns generated by the zigzag motion of the GPR and creating a 3-D image that could be easily interpreted. Once the pattern recognition for pipe like structures has been studied, the next stage would be mounting the GPR with the EMI and obtaining a better estimation of the presence or absence of buried pipes. A sample GPR image of a buried pipe is shown in Figure 5.

5.INCREASING EFFICACY OF BUDS USING MULTI-SENSOR APPROACH

As mentioned above, the aim is to create a real-time, accurate, and most importantly reliable fish-finder type utility detection system with minimal false positives. Although the two sensors discussed above have their advantages, alone neither can give reliable estimation of the buried utilities. For example, the GPR module is reasonably accurate on the depth of the object, but is unable to distinguish between a metallic and a plastic pipe. Similarly, the EMI, by itself is not capable of providing accurate information regarding the position (e.g. depth) of a metallic object. These limitations of the individual sensors can be overcome by a fusion of sensor data.

According to Klein (1993), data fusion is a multilevel, multifaceted process dealing with the automatic detection, association, correlation, estimation, and combination of data and information from multiple sources. The type of fusion architecture used to combine sensor data depends on the application. There are three broad 'levels' where data fusion can be incorporated: a) direct fusion of sensor data, b) feature vectors and c) high-level inferences. Since the multi-sensor data in our case is not commensurate, we can either represent data obtained from each sensor via feature vectors, with their subsequent fusion; or perform individual processing of each sensor's data to achieve independent high-level inferences or decisions, which are combined to make a collective decision.

The overall process consists of four main parts: Preprocessing the signal from the sensor, feature/contour extraction, feature/property selection, and classification. Figure 6 shows the adapted version of a common data fusion model to MS-BUDS (Multi-Sensory Buried Utility Detection System). The data from the two sensors are initially conditioned by independent signal processing modules which later feed into a parallel processor running a multisensor algorithm. The parallel processor uses a feature recognition and classification algorithm that operates on sensory data and 'learns' with a knowledge base as it goes along.

6. SUMMARY

Damage to buried utilities can cost lives and damage to property and equipment. This paper presented a novel technology that integrates two common sensory equipment, the Pulse EMI and GPR, into a multi-sensory real time underground utility detection system. The premise of fusing the two sensory data stream is to maximize reliability/accuracy while minimizing false positives. Two experimental facilities have been built to study the effect of various soil and object conditions on the features of the sensory outputs. In the next phase two experimental facilities will be combined into one platform, the MS-BUDS.

7. ACKNOWLEDGEMENTS

Funding for the work has been provided by Public Health Service-National Institutes of Health under contract 1R01 0H04201-01.

8. REFERENCES

Das, Y., McFee, J., Toews J., Stuart G.: "Analysis of an Electromagnetic Induction Detector for Real-Time Location of Buried Objects", IEEE Transactions on GeoScience and Remote Sensing v.28 No.3, May 1990.

Gader, P., Mystkowski, M., Zhao, Y.: "Landmine detection with GPR using Hidden Markov Models", IEEE Transactions on GeoScience and Remote Sensing v. 39 No. 6, 2001.

Klein, L.: "Sensor and Data Fusion Concepts and Applications", v. TT14, SPIE Press, 1993.

Lorenc, S.J., Bernold, L.E: "Excavator-mounted Ordnance Locating System for Remote Excavation", 14th International Symposium on Automation in Construction, June 1997.

Underground Focus Magazine, (1999). Canterbury Communications, Spooner, WI. URL: http://www.underspace.com/

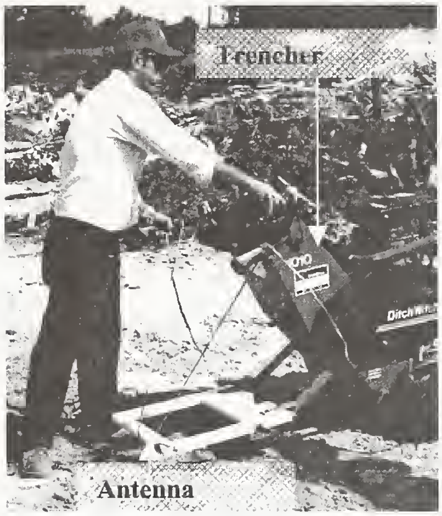


Fig. 1. Equipment mounted BUDS

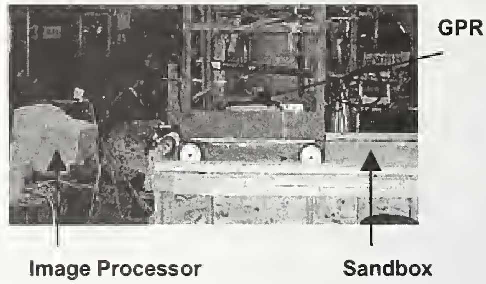


Fig 4. GPR setup for utility detection.

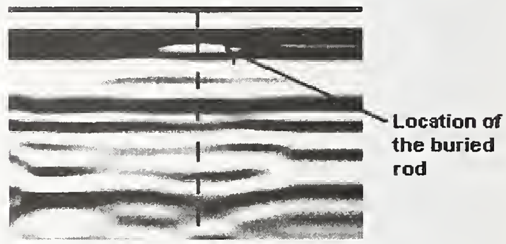


Fig 5. Sample GPR image of a buried rod

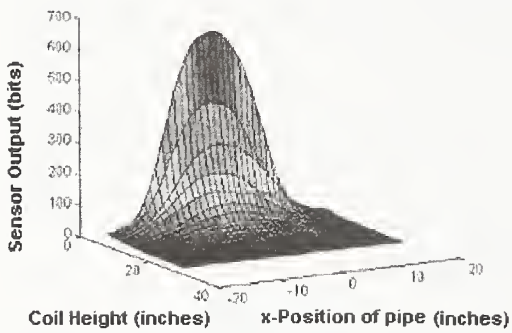


Fig 2. Graphical sensory output of BUDS

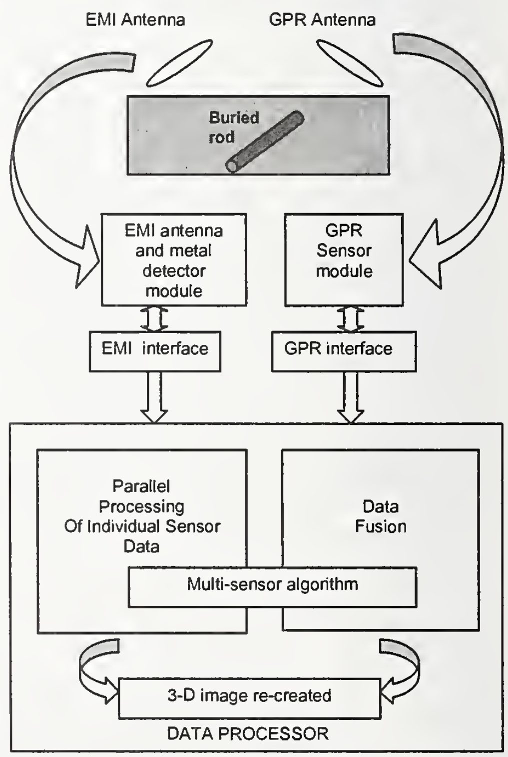


Fig 6. Simplified data fusion model adapted to MS-BUDS

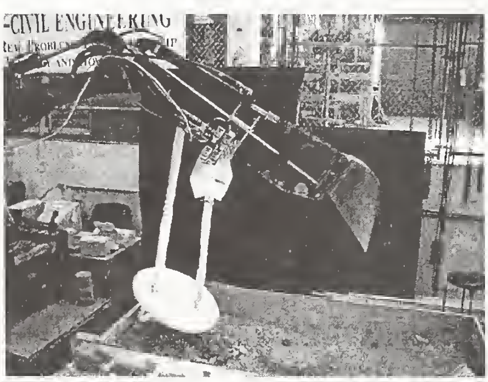


Fig 3. Buried Utility Detection System mounted on a backhoe.

Automated Inspection of Utility Pipes: A Solution Strategy for Data Management

Tariq Shehab-Eldeen[1] **and Osama Moselhi**[2]

(1) Post-Doctoral Fellow, Department of Building, Civil and Environmental Engineering, Concordia University, Canada.
tshld@cbs-engr.concordia.ca.

(2) Professor and Chair, Department of Building, Civil and Environmental Engineering, Concordia University, Canada.
moselhi@cbs-engr.concordia.ca

ABSTRACT: Rehabilitation of urban infrastructures has received considerable attention in North America, creating a need for automation. Automating the rehabilitation process of various infrastructure facilities is driven by the need for cost reduction, higher quality and improved safety. This paper describes an automated system, AUTO-DETECT, recently developed, for rehabilitation of sewer pipes. AUTO-DETECT automatically analyzes the CCTV videotapes that depict the conditions of the surveyed pipes and consequently detects and classifies defects. It introduces five sets of specialized neural networks, each is dedicated for one type of defect. The paper also presents the integration aspects of these five sets of neural networks to formulate a solution strategy that is utilized to improve the performance of the developed diagnostic system.

KEYWORDS: Automation; Sewer pipes; Inspection; Data management.

1. INTRODUCTION

Rehabilitation of urban infrastructures has received considerable attention in North America, creating a need for automation. Automating the rehabilitation process of various infrastructure facilities is driven by the need for cost reduction, higher quality and improved safety. A typical maintenance or rehabilitation process of underground sewer pipes usually starts by surveying these pipes and collecting relevant data about their condition. This data usually highlights many aspects and provides useful information such as the presence, type, number and location of defects. CCTV (closed circuit television) cameras are commonly used to capture these data. CCTV cameras produce a videotape that has to be visually inspected by a human expert in order to identify and locate defects, if they exist. The process is usually time consuming, tedious and expensive. Interviews conducted with several municipal engineers and consultants in Quebec and Ontario, Canada revealed that the cost of sewer inspection is about CDN $1.5 per linear meter, 30 % of this cost (i.e. $0.42) is spent on inspection of videotapes (Shehab-Eldeen 2001).

This paper describes an automated system, AUTO-DETECT, recently developed, for rehabilitation of sewer pipes (Shehab-Eldeen 2001). AUTO-DETECT automatically analyzes the CCTV videotapes that depict the conditions of the surveyed pipes and consequently detects and classifies defects. It utilizes image analysis techniques and artificial intelligence (AI) to perform its task through five sets of specialized neural networks, each set consists of three networks. Unlike the work developed earlier by the authors (Moselhi and Shehab-Eldeen 2000) where one classifier was developed to detect different types of defects, this paper introduces five sets of specialized neural networks, each is dedicated for one type of defect. This paper also presents the integration aspects of these five sets to formulate a solution strategy that employs a multiple classifier technology, designed to improve the performance of the developed system. An example application is presented to demonstrate the use and capabilities of the developed system.

2. DEVELOPED SYSTEM

The developed system makes use of and builds on current practice. The process used in current practice for detecting defects has been described elsewhere (Moselhi and Shehab-Eldeen 1999 (a) and (b)). Figure 1 depicts the overall configuration of AUTO-DETECT. As shown in this figure, a closed circuit television (CCTV), or a zooming, camera first scans the inner surface of a pipe and produces a videotape which is then played back using a VCR. The VCR then feeds the information captured on the tape to a computer equipped with a frame grabber and multiple classifier modules. The frame grabber captures and digitizes the frames of the acquired images. The multiple classifier module utilizes an image analysis software package to analyze those digitized frames and processes them in a manner so as to prepare a suitable input (i.e. feature vectors) to each classifier. A solution strategy is designed to integrate these classifiers (See Figure 2). The feature vectors are then fed into the developed system and are accordingly classified into five categories of defects. These categories are deposits, joint displacements, cross-sectional reductions, infiltration and cracks. This paper focuses primarily on the solution strategy module, other modules have been presented elsewhere (Moselhi and Shehab-Eldeen 2000 and 2001).

3. SOLUTION STRATEGY

Neural networks work in an analogous way to human experts. The more focused, i.e. domain specific, they are, the higher their problem solving capabilities. In order to express and demonstrate the importance of specialty in classification tasks, several classifiers (i.e. neural networks) were developed; each is considered suitable for a certain category of defects. This was considered advantageous, as opposed to one network that classifies more than one type of defect. Although diversity of networks is advantageous, it may lead to a problem in guiding the detected patterns to the proper channel so as to ensure that each category of defect is received by the most suitable specialized classifier. To overcome this problem a solution strategy is presented to organize data traffic so as to guide the patterns in an efficient manner and accordingly improve the performance of the system.

Figure 2 depicts the proposed solution strategy. As shown in this figure, all images are processed three times. In the first pass (i.e. inverted images), all images are inverted, dilated, background subtracted, thresholded, segmented and finally analyzed. In the second pass (i.e. non-edge detection), images are subjected to the same image processing techniques except for inversion. In the third pass (i.e. edge detection), all images are subjected to a number of operations such as background subtraction, edge detection, dilation, thresholding and analysis. The reason behind subjecting the same videotape to a number of passes is to benefit from all image processing techniques that are necessary to detect all categories of defects recognizable by the system.

As can be seen in Figure 2, patterns depicted on images subjected to the first pass (i.e. inverted images) will first be processed by set of networks number 1, specialized in classifying deposits. These networks will classify the input data (i.e. patterns) into two categories: “Deposits” and “Else” (i.e. non-deposits). All patterns classified as “Else” will be screened based on their X and Y coordinate and will be further processed by another two sets of networks (i.e. sets no.2 and 3), each is specialized to deal with a specific set of defects. Patterns with X and Y coordinates located only at the center of an image will be fed into the networks specialized in classifying cross-sectional reductions and misalignments (i.e. set no. 2 and 3, respectively). Patterns classified as “Else” by set # 2 and 3 will be ignored being either non defects or defects that are not recognizable by the system. It should be noted that the system recognizes more than 90% of defects that commonly exist in sewer pipes (Moselhi and Shehab-Eldeen 1999(b)).

Patterns depicted on images subjected to the second pass (i.e. non-edge detection) will be fed into the networks specialized in classifying infiltration (i.e. set no 4). These networks are capable of classifying patterns into two categories: “Infiltration” and “Else” (i.e. non-infiltration). It should be noted that all patterns classified as “Else” are considered as either non defects or defects that are not recognizable by the system.

Patterns depicted on images subjected to the third pass (i.e. edge detection) will be fed into the networks specialized in classifying cracks (i.e. set no 5). These networks are capable of classifying patterns into two categories: "Crack" and "Else" (i.e. non-crack). It should be noted that all patterns classified as "Else" are considered either non defects or defects that are not recognizable by the system. It should be noted that each of the five sets consists of three neural networks (Moselhi and Shehab-Eldeen 2001 and Shehab-Eldeen 2001).

4. EXAMPLE APPLICATION

To demonstrate the use of the developed system and the capabilities of its solution strategy module, the image shown in Figure 3 was considered. It should be noted that due to space limitations, the case example will focus primarily on the third pass (i.e. detection and classification of deposits using edge detection).

As can be seen, the image depicts a number of objects. These objects are cracks and a number of non-defects. To detect and classify these objects, the image was processed in the same manner as explained earlier. The image was segmented as shown in Figures 4. As can be noticed 15 objects were detected. The feature vectors describing these objects were then fed into two classifiers. The first is specialized in cracks, while the second was trained to classify four types of defects: (1) cracks; (2) multiple cracks; (3) cross-sectional reductions and (4) misalignments. The results obtained from the specialized and non-specialized classifiers are shown in Figure 5 and 6, respectively. As can be noticed that the specialized classifier reduced the false alarm for presence of cracks by 50%. Clearly this finding, while indicative of benefits of multiple specialized classifiers, can not be generalized.

5. CONCLUSION

An automated system for detection and classification of defects in sewer pipes has been presented. The system utilizes image analysis, solution strategy and multiple classifier modules to performing its task. To demonstrate the importance of specialty in classification tasks, several classifiers (i.e. neural networks) were used; each is considered suitable for a certain category of defects. These classifiers are specialized in deposits, cross-sectional reductions, misalignments, infiltration and cracks. The paper focused primarily on presenting a solution strategy that was developed to organize data traffic so as to guide the extracted feature vectors (i.e. signatures) of various defects to a set of specialized neural networks in an efficient manner. This was carried out in order to improve the overall performance of the developed system. This was considered advantageous, as opposed to one network that classifies more than one type of defect. A case example was also presented.

6. REFERENCES

Moselhi, O. and Shehab-Eldeen, T. (1999) (a),"Automated Detection of Defects in Underground Sewer and Water Pipes." Journal of Automation in Construction, Elsevier Science, 8, 581-588.

Moselhi, O. and Shehab-Eldeen, T. (1999) (b)." An AI-Based System for Detection and Classification of Defects in Sewers." Proceedings of INFRA 99 International Conference, Center for Expertise and Research on Infrastructures in Urban Areas (CERIU), Montreal, Canada, 3B: 42-54.

Moselhi, O. and Shehab-Eldeen, T. (2000). " Classification of Defects in Sewer Pipes Using Neural Networks." Journal of Infrastructure Systems, ASCE, 6 (3), 97-105.

Moselhi, O. and Shehab-Eldeen, T. (2001)." Multiple Classifiers for Automated Detection of Defects in Sewer Pipes"." Proceeding of 2001 International Conference on Rehabilitation of Infrastructures, Waterloo, Canada, 273-278.

Shehab-Eldeen, T. (2001)." An Automated System for Detection, Classification and Rehabilitation of Defects in Sewer Pipes." PhD thesis, Concordia university, Canada.

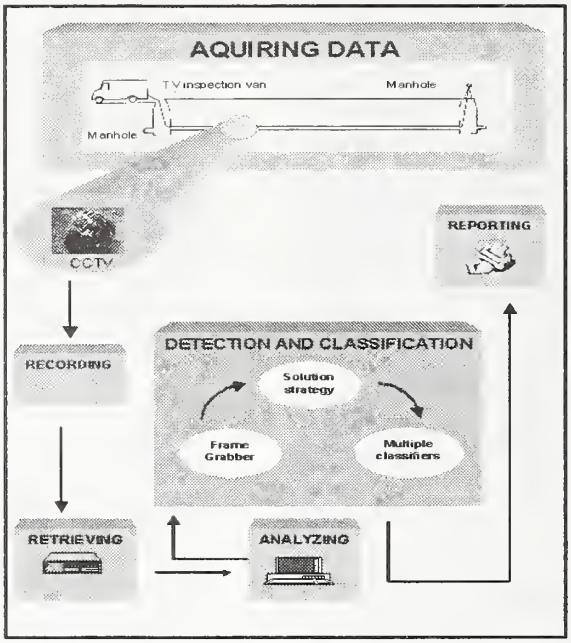


Figure 1: Developed Automated Detection and Classification System

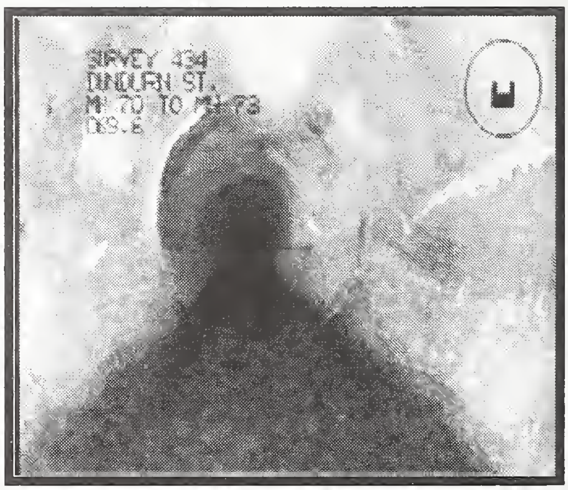


Figure 3: Original Image

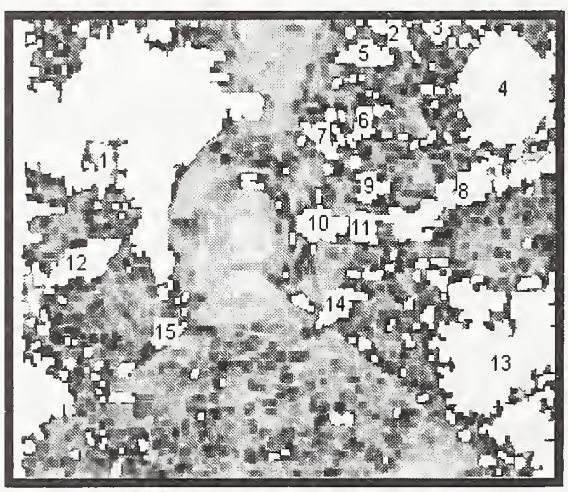


Figure 4: Segmented Image

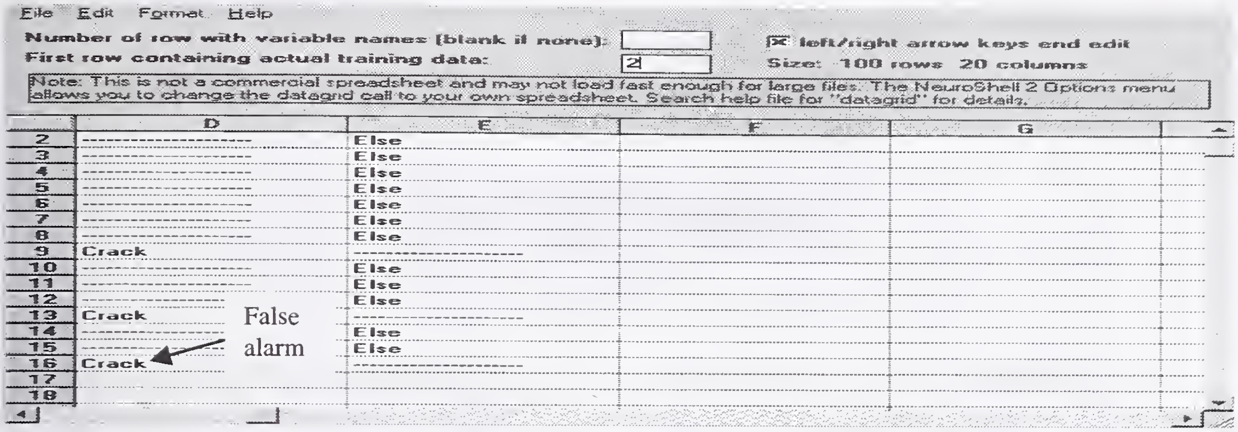


Figure 5: Results Obtained Form the Specialized Classifier

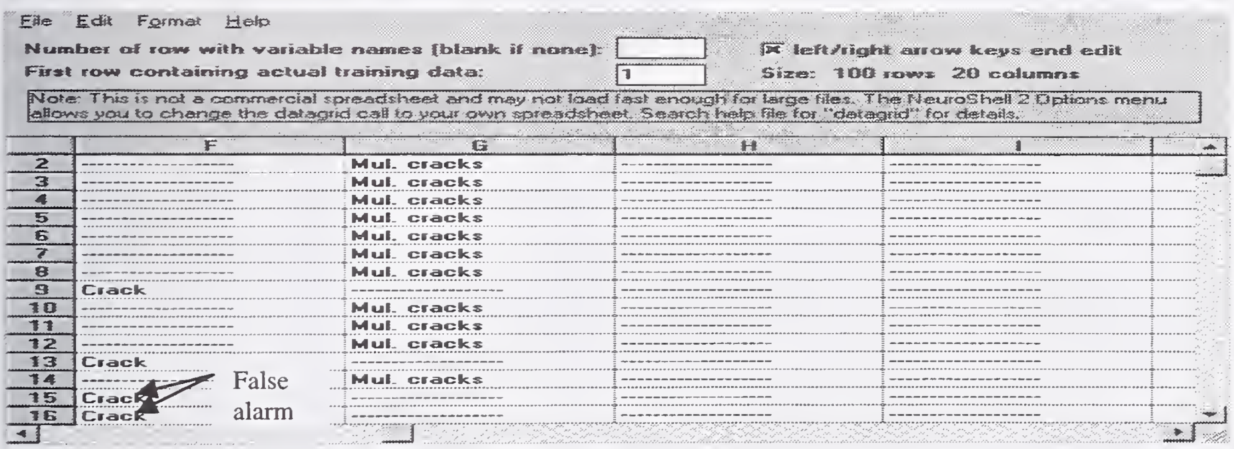


Figure 6: Results Obtained Form the Non-specialized Classifier

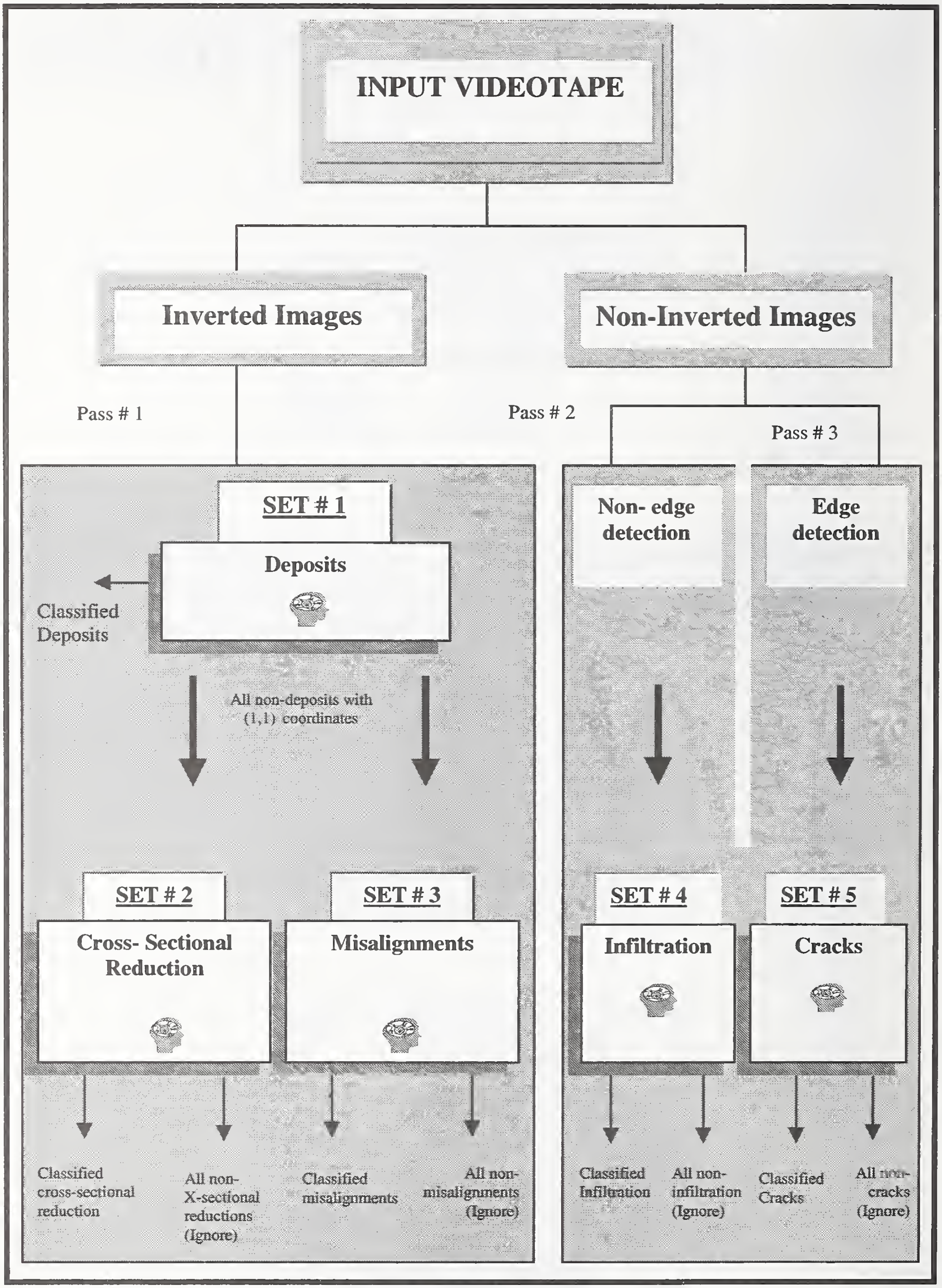


Figure 2: Solution Strategy

A Concept of the Robotoid Manager with AR

Thomas Bock, Shigeki Ashida

The Chair for Realization and Informatics of Construction
The Faculty of Architecture, Technical University of Munich
Arcisstr. 21 80333 Munich, Germany, Tel. +49-89-289-22100, Fax. +49-89-289-22102
Email: Thomas.bock@bri.ar.tum.de : ashida@bri.ar.tum.de

ABSTRACT: For efficiency, construction managers take care of several job sites simultaneously. In order to be omnipresent, we propose a notification of ubiquitous construction management.
Here we try to adopt the system, so named the Robotoid Manager, with;

- AR system with mobile computers
- Message board information system
- Multi schedule optimizing
- Progress monitoring and schedule modification

In addition, we have taken an embedded system into account for a fast and easy recognition of real situation. Total efficiency of management of several sites is improved.

KEYWORDS: Ubiquitous management, AR (Augmented reality), Message board, Schedule optimizing, Schedule modification, Embedded system, RFID (Radio Frequency Identification)

1. INTRODUCTION

For efficiency, construction managers recently take care of several sites simultaneously in many cases. In order to be omnipresent, we propose a notification of ubiquitous construction management by adopting of augmented reality system with mobile computers.

It is able for a robotoid manager to reduce a time of transfer to each job-site, and to have communication better with more people who are not on the site and even not concerned directly with the construction. Just-in-time decision and fast information transmitting will be possible, because of accessing updated and large volume of data from various aspects, whenever and wherever he/she is.

There are also a lot of chances to do another task at a different situation while managing. It reduces a total required time of work.

2. Difficulties

2.1. Differences of situation

Under a ubiquitous management, the situation of managing may often be quite different from the site. There are also many other disturbances. For example, there are many people, e.g. in an office, who are not concerned on the site.

And processes on the site are not recognized directly. If the manager is always at the site, he/she can supervise all activities. Sometimes it is said that these human senses are most important to make a best decision. A robotoid manager must recover this lack of recognition as well by using additional communication devices.

2.2. Confusion of several sites

Different situations of each site may be sometimes not allocated to the proper site, especially caused by unexpected apprehension or unclear information. The manager must always take it into account, which information belongs to which site. Misunderstandings may sometimes result in a serious loss of efficiency.

3. A concept of the Robotoid Manager

3.1. Possibilities of AR system

Here we can adopt an augmented reality system (shortly AR) with mobile computers for a robotoid manager. Many of these systems have been already introduced in the market; sometimes as wearable style that means it is easy to carry and use. It provides anywhere a similar vision as that on the site, and also free from a disturbance in different surrounding conditions. A robotoid manager can always wear all related data and communication tool with visual help of AR. AR can also increase efficiency of communication without oral method.
With an AR system, in addition we can take following possibilities.

- Visual comparison between a plan and a real progression.
- High visibility
  Without a dead angle / Through obstacles / Inner-structure / Underground / Clear through dense dusts, darkness, mud, water etc.
- Visual help for a consideration
  Situation of temperature, Humidity, air flow, direction and level of lighting, smell, sound and others
- Indication for a pointing of position and time
- Management of construction materials and parts
- Pre-check of a process and effect simulation
- Communication crossover situations
- For training

In case information should be shared by other people, it can either be projected on screen or on each person goggles. For better visibility, a larger screen is necessary, which is harder or more troublesome to carry. Then a useful indication device, for example a mobile or wearable beamer, should be the best combination.

3.2. Message board information system

A manager must always be aware of a present situation, which changes time to time, in order to make a just-in-time decision. Therefore it is better to use real time data and their recent previous related actions, instead of using too much stored data. Too much data require many capacity of computer server, and they are sometimes never used anymore. And also it requires for a manager complicate searching, comparing and selecting the proper data.

To get real new data while classifying, message board information system (Fig.1) is a practical method. This method means to gather opinions of all workers at each site at the real time. Due to the use of this system, we can spare sub managers on various site. An Internet message board system with a mobile communication device has been already used. So almost all worker on site can use this system easily without any complicate installation or any expensive devices.

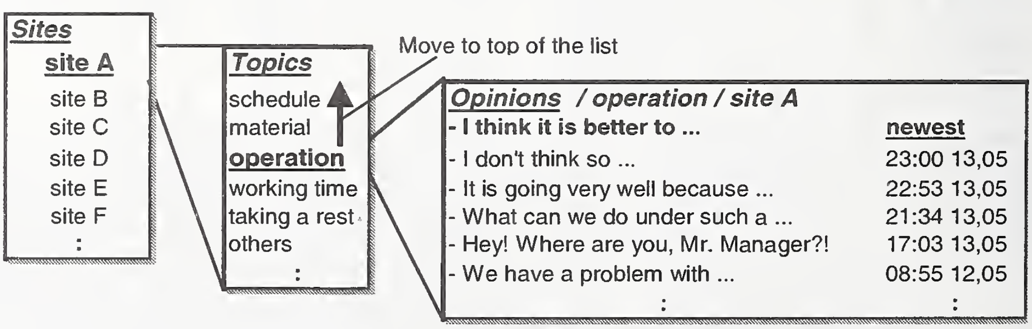


Fig.1. Message board information system

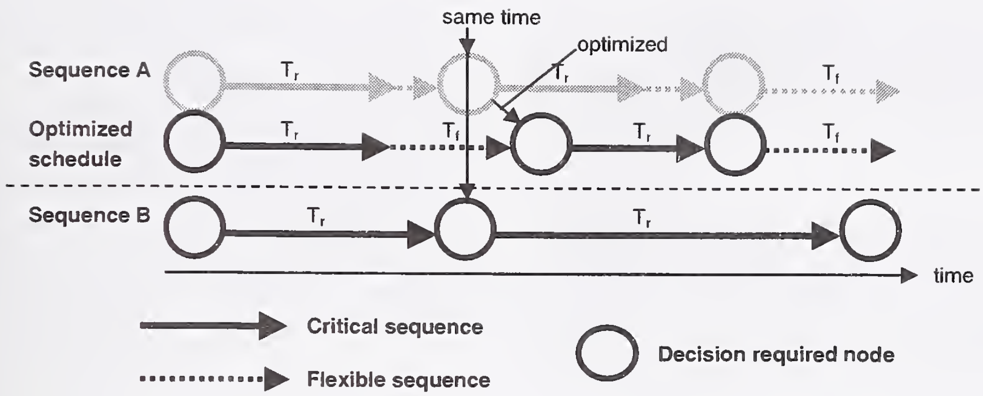


Fig.2. A concept of multi schedule optimizing

Each information is already classified by its site and topic (interest). Always newest information is listed up at the top. Opinions of both "goes good" and "problem, decision required" can be collected at the real time. Special topics, in which the workers on site are actually interested, occur naturally in the message board by themselves. The more they are interested in a topic, the more arguments about the topic are listed. This is an indication of the necessity for a manager to take action or to be alert.

This system is effective not only collecting opinions, but also informing a decision of the manager.

3.3. Optimizing the combination of each other site

While sequences on several sites are progressing in parallel, and when some of them are a decision required at one time, the rest of them must be wait for a decision. So it is the best that there comes always one decision required sequence after another. A combination between automated sequences and decision required sequences should be optimized automatically with a multi schedule optimizing method.

All sequences are related each other owing to its order and required time. It is indicated as an arrow diagram. A progress is monitored and its criticality of time is considered. A flexibility of the sequence is evaluated as follows;

$$T_f = (t_l - t_e) - T_r \qquad (1)$$

Here, T_f : a flexible time, t_l : the latest time to be allowed, t_e : the earliest time to be able to start, T_r : a preset required time

When $T_f=0$, then the sequence is critical and the decision for the next sequence must be prepared before the time of
t_l in order not to make delay of total construction period. A concept image of multi schedule optimizing is shown as a part of arrow diagram in Fig. 2.

And the optimized schedules are always possible to be indicated on this managing system.

3.4. Modification of a preset schedule

In case of an unpredicted disturbance or change of the condition, the manager must realize its pre-signal and make a better decision as soon as possible.

For a changed situation, the schedule of each site must always be modified with an analysis of the trend. We can take t_e into account at this modification as follows;

$$t_e(n) = t_e(n-1) + T_p (+ T_u) \qquad (2)$$

Here, $t_e(n)$: t_e of the sequence "n", $t_e(n-1)$: t_e of the previous sequence of "n", T_p : a required time which is predicted from a real progress of the sequence, T_u : a required time of the unscheduled sequence (if necessary)

The index t_e is always rewrote owing to a monitoring of the progress. Rewriting concept is shown in Fig.3. T_p is predicted by comparing a progress with a preset required time. There may be also a case that an unscheduled sequence becomes necessary.

And the optimizing of a combination, which is described before, is repeated over and over automatically for all these modification to make the total efficiency higher. This system indicates all better possibilities up to this optimizing whenever a decision is required for a ubiquitous manager.

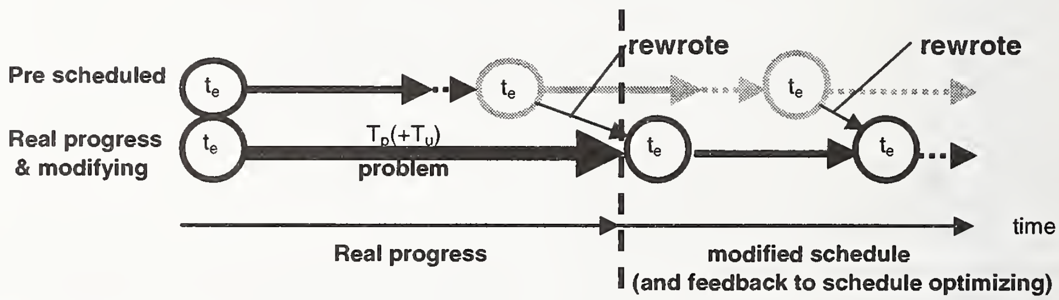


Fig. 3. Rewriting concept on the modification

4. Situation recognizing with embedded system

4.1. Installation of RFID

All decisions should be always made regarding a present requirement. Making a decision effectively we analyse the present situation and refer the consolidated findings in the future.

3D scanning with AR is one of the best methods to grasp the real situation. AR itself with wearable computers has also already been popular. But in the use of AR for 3D scanning, there are still some technical and practical difficulties to realize. It is sometimes expensive, inexact and time for scanning is required. Also it is still hard to carry or wear the equipment easily, for example still many devices are required as that for recognizing self position. Therefore an accepted technique method only with visual 3D scanning for every concerned site is not practice to use, at the moment.

On the other hand, RFID (Radio Frequency Identification) is one of the popular identification methods. The chip is embedded in all objects at each site containing its specific data. The scanner for this system is in most of the cases easy to carry or wear.

4.2. Concept of recognizing

We are here trying to adopt an embedded system with RFID and AR system with mobile computers for recognizing of sites to a robotoid manager.
Specific data of all objects at the site are preliminary registered and embedded with RFID chip. In case, more detail data must be stored, the identification is related to data bank automatically (Fig. 4).

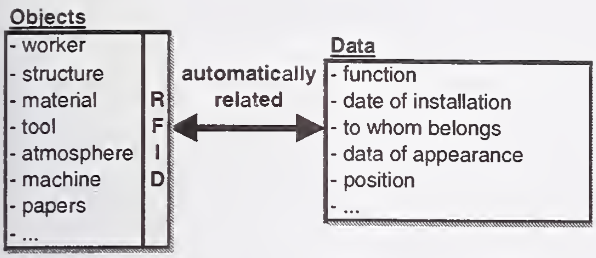


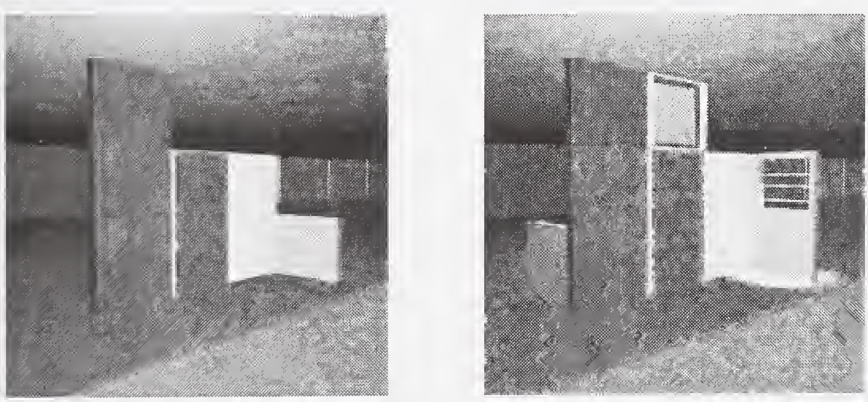
Fig. 4. Embedded data

With RFID all objects on the site in the radio effective distance are identified at the same time and grasped its specifications fast and easily. The visual situation, which means position, wearing, movement and others, is shown virtually with AR, like as Fig. 5. The manager can take also an invisible object into account. The ability of RFID for rewriting is important as well. On the view of AR, a robotoid manager compares the real present situation with the virtual view, which made from pre registered RFID data (Fig. 5). Then the difference is rewritten on RFID system.

5. Future work

This trial is still just a concept. There has been no real adoption yet. We must think furthermore about the realization. The concrete systems should be created and combined at its best. And of course its efficiency must be confirmed with some case study or practical adoption.

Furthermore this method will be used in many other kind of management, for example buildings, facilities, infrastructures and others. Therefore this method should be not only used to analyse the construction site, but also included into total system of management to be used effectively. We will try to implement such a robotoid management system with AR.



a) real situation **b) scheduled situation**

Fig.5. Comparison on the view of AR

Representation and Integration of As-Built Information to IFC Based Product and Process Models for Automated Assessment of As-Built Conditions

Burcu Akinci[1], Frank Boukamp[2]

ABSTRACT: Frequent assessment of as-built conditions is necessary for active project control during construction. Assessment of as-built conditions involves collecting accurate as-built information and comparing the as-built information collected with given design and schedule information. Currently, different technologies exist for collecting accurate and comprehensive as-built information. The as-built information generated through these technologies is then manually evaluated by comparing it to the design information. This process is time-consuming and error prone. Therefore, there is a need for an automated assessment of as-built conditions. This paper discusses the applicability of the Industry Foundation Classes (IFC) in their current version 2x to support the assessment of as-built conditions of construction projects. The goal is to create a project model that represents and integrates as-built and design information. In this paper, we state different types of information that need to be represented in this integrated model, describe common ways of representing this information in IFC Rel.2x, and highlight the limitations of current IFC Rel.2x specifications in representing an integrated design and as-built model. The paper concludes with a proposal for extending the IFC specifications to enable the automated assessment of as-built conditions.

KEYWORDS: Industry Foundation Classes, integrated design and as-built models, construction

1. Introduction

Frequent and comprehensive assessment of as-built conditions is necessary for project control and to minimize delays caused by late detection of defects at construction sites. A comprehensive assessment of as-built conditions involves frequent, complete and accurate collection and storing of data on as-built conditions, and a formalized approach for comparing the as-built conditions with the as-designed requirements.

Different technologies can now enable the frequent, complete and accurate collection of as-built conditions. For example, laser-scanning technologies are gaining acceptance in the A/E/C (Architecture/ Engineering/Construction) industry because of their ability to create complete and accurate 3D as-built environment models based on spatial information [1]. However, currently, even if a laser scanner is used at a construction site, the comparison of as-built models with existing design information is still being performed manually by visually inspecting both design and as-built information according to a construction schedule and design specifications [2]. This process is time-consuming and error prone. Therefore, there is a need for an automated assessment of as-built conditions.

To automate the assessment of as-built conditions, the as-built information, consisting of as-built product model and schedule, needs to be integrated with the as-designed product and process information. Both as-built and as-designed models need to be represented in a semantically rich way and the necessary relationships between these two models need to be created and maintained throughout construction.

This paper discusses possible ways to integrate as-built information within IFC (Industry Foundation Classes) based product and process models, and evaluates the IFC Rel.2x specifications in terms of its ability to represent integrated design and as-built information. The discussions in this paper

[1] Assist. Prof., Dept. of Civil and Environ. Eng., Carnegie Mellon University, Pittsburgh, PA 15213, bakinci@cmu.edu

[2] Research Assist., Dept. of Civil and Environ. Eng., Carnegie Mellon University, Pittsburgh, PA 15213, boukamp@cmu.edu

are based on the implementation of a small portion of a warehouse construction case.

2. Background

Currently different industry efforts exist for modeling and merging design, construction and FM (facility management) information. For example, the Steel Construction Institute is developing CIMsteel Integration Standards (CIS), a framework for modeling construction projects that are mainly based on steel structures. The goal of this effort is to "allow software vendors to make their engineering applications mutually compatible" [3]. As another example, the International Alliance for Interoperability (IAI) is developing the Industry Foundation Classes (IFC), a similar framework that shall allow making A/E/C and FM related information exchangeable between different software programs by standardizing the representation of A/E/C and FM related information within project model descriptions. In our research, we decided to use the IFC representation since it does not focus on only one specific industry and hence it is more generally applicable.

Currently IFC models are either used to store design or as-built information. Often an IFC-based design model is created and later on updated with as-built information, overwriting the designed information. To allow the storage of different versions of a project, like 'Schematic Design', 'Detailed Design', 'As-Built', etc. these models are usually stored into different files, which then are associated with the description of the version. For example Metracker, a performance metric tracking tool, uses this approach [4].

To be able to automate the assessment of as-built conditions efficiently, we need to represent detailed design information and as-built information in one IFC model. An integrated representation of design and as-built models allows to identify and reason about the relationships between design model and as-built model efficiently, and hence minimizes the computational efforts associated with checking one model with the other to assess as-built conditions.

Section 4 of this paper describes how we build on and extend the current IFC Rel.2x specifications to develop an integrated design and as-built model for the case that we implemented.

3. Case Study Description

In November 2001, we performed a case study at a warehouse construction project in Pittsburgh, PA, where we collected design and schedule information through conventional methods and as-built information through laser scanners. The main structure of the warehouse consists of steel frames. We modeled the design and as-built information of one of the steel frames using IFC Rel.2x to automate the assessment of as-built conditions for that frame.

Assessment of as-built conditions in the frame example involves comparing the as-designed and the as-built geometric, material and schedule information. Figure 1 shows the design and as-built information of the frame that we modeled. In this case, the contractor used the right material and built it in the right location. However, the as-built schedule deviated from the as-designed. Hence, an automated assessment system should be able to compare the different types of information represented in as-design and as-built model and identify the schedule deviation in this case.

4. Integrating Design and as-built Information Using IFC

Since the assessment of as-built conditions requires reasoning about material information, building element type information, geometric information, scheduling information and relationships between building elements, we need to explicitly represent this information. In IFC Rel.2x it is usually represented as follows:

Material information is usually assigned to the respective *IfcObject*-object, through a relationship called *IfcRelAssociatesMaterial* (Figure 2).

BUILDING ELEMENT TYPE INFORMATION is usually modeled through properties or type libraries defined by relationships of type *IfcRelDefinesByProperties* (Figure 3) or *IfcRelDefinesByType* (Figure 4) respectively.

GEOMETRIC INFORMATION is modeled by grouping *IfcRepresentation*-objects in an *IfcProductRepresentation*-object that is then assigned to a building object (Figure 5) or through properties providing geometric information being assigned to an *IfcObject*-object by an *IfcRelDefinesByProperties* relationship (Figure 6).

SCHEDULING INFORMATION is represented in IFC by assigning an *IfcTask*-object, representing a task in a schedule, to an *IfcObject*-object, via a relationship of type *IfcRelAssignsToProcess* (Figure 7).

CONNECTIONS BETWEEN BUILDING ELEMENTS are modeled by using *IfcRelConnectsElements* relationships.

As stated before, IFCs were developed to allow the representation of either design or as-built information in an interoperable product model. When trying to incorporate both into one model, three scenarios might be observed [5]:

1. The current IFC model representations should be used if possible.
2. The current IFC model representations can be extended if necessary.
3. New concepts are developed only if a sufficient representation cannot be created using scenarios 1 and 2.

Since minimal changes to the IFC model are desirable, first we tried to use existing representations to integrate the as-built information into the IFC-based model and then suggested extension to the current representations before we develop any new representation schema.

4.1 Utilization of current IFC Rel.2x Representations

Three approaches seemed promising in representing integrated as-designed and as-built information using the current IFC Rel.2x specifications:

(1) Utilization of representation context concept defined through *IfcRepresentationContext*.

IfcRepresentationContext objects, to which product representations can be assigned, allow products to have different representations, e.g. one that represents a sketch of the product and one that represents the detailed design of the product. To do this, the C*ontextType* attribute of this object is being proposed to be used to store different context information such as 'Sketch' and 'Design' to distinguish between the different contexts. By using 'Design' and 'AsBuilt' as values for this attribute we would be able to create two different representation contexts for a building element to which we can assign design and as-built information respectively. This would allow having the design and as-built information incorporated into one IFC-based product model. Currently, only *IfcRepresentation* objects can be assigned to an *IfcRepresentationContext*. As a result, only geometric or topologic information can be assigned to these different contexts. However, as we discussed in the case study, other information, such as scheduling information, will also have to be kept in distinguishable design and as-built contexts. The usage of different *IfcRepresentationContexts* thus becomes insufficient for our purposes.

(2) Utilization of logical grouping concept to group as-designed and as-built information.

IAI provides the possibility to group information by using *IfcGroup*-objects. Based on this, we can group design information and as-built information in a design-group or an as-built-group through the assignment of the information to the respective *IfcGroup*-object (Figure 8). The design and as-built information can be assigned to the respective groups by using *IfcRelAssignsToGroup* relationships. The *IfcRelAssignsToGroup*-relationship is a subtype of *IfcRelAssigns*-relationships, which can only link subtypes of *IfcObject* among each other. This creates a problem for representing designed and as-built material information using *IfcMaterialProperties*-objects which are not derived from *IfcObject*. Since it is important to represent and reason about the material information in design and as-built models, the *IfcGroup*-based approach becomes inadequate for our purposes.

(3) Creation of new properties of entities using *IfcProperty* object.

This would involve the creation of *IfcProperty*-objects having a specific value in its *Name*-attribute which tells whether it describes design or as-built information. These *IfcProperty*-objects then will be assigned to the related entities through an *IfcRelDefinesByProperties*-relationship. The usage of properties for extending IFC concepts is the most common technique used. Currently properties can only be assigned to objects derived from the class *IfcObject*. This again creates a limitation of the usability of this concept for our purposes, since in IFC for example material information or relationships in general are not derived from *IfcObject*.

4.2. Extended Concepts

Since the current IFC Rel.2x specifications do not allow us to model the as-built conditions completely we investigated other ways to incorporate design and as-built information into one project model with minimal extension of existing IFC concepts.

Extending the concept of reified relationships looked most promising for our purposes. Most of the information that we need to model to support automated assessment of as-built conditions is assigned through relationships to the objects (Figures 2-4, 6, 7). All of these relationships are represented as instances of classes that are derived from the superclass *IfcRelationship*. Thus, if we add a new attribute called *Context* to the *IfcRelationship*-class, then we can use that attribute to determine whether the information is related to the design or the as-built context. This *Context*-attribute should be of type *IfcLabel* and have a descriptive name, e.g. 'Design' or 'AsBuilt', assigned as value.

To enable the interoperability of the model, it is necessary to standardize these values that will be assigned to the context-distinguishing attributes. Otherwise, software developers using IFCs will have to agree on how to describe the different contexts to be able to interpret the models that are interchanged using their software.

Enabling the class *IfcRelationship* to describe the context it is assigned to through the new *Context*-attribute allows us to keep nearly all the information that we need for automated defect detection in separate contexts, i.e. a design and an as-built context. This is because all the information is assigned through subtypes of *IfcRelationship* to the entities in the IFC model (Figures 2-4, 6, 7). With this extension we can provide a 'Design'-relationship to the design information of an element and an 'AsBuilt'-relationship to the respective as-built information of an element.

However, this extension of the IFC concept of objectified relationships is not entirely sufficient, since the geometric information is the only information that may not be assigned to an object through a relationship but may also be described through *IfcRepresentation*-objects that are directly assigned to the element not using *IfcRelationship*-objects.

4.3. Proposed Concept

In the previous two sections, we described four different approaches to incorporate design and as-built information into one integrated model:

- using *IfcRepresentationContext*-objects

- using *IfcGroup*-objects
- using new *IfcProperty*-objects
- using an extended version of *IfcRelationship*

All of these approaches, when used alone, are insufficient for the purpose of automated as-built assessment. Instead of creating a completely new concept for the integration of design and as-built information into one IFC-based project model, we propose combining two of the approaches described above as the proposed way of modeling as-designed and as-built information in an integrated way.

Our proposed approach combines utilization of different *IfcRepresentationContext*-objects and extending the class *IfcRelationship* as explained before (see Figure 9).

The weakness of the concept of extending the class *IfcRelationship* was that the part of the geometric information that is described through *IfcRepresentation*-objects cannot be distinguished by using 'Design'- and 'AsBuilt'-relationships. This weakness can be overcome by the usage of different *IfcRepresentationContext*-objects. The *IfcRepresentation*-objects, representing the geometric information of the entities in the IFC model, are assigned to a 'Designed'- or 'AsBuilt'-representation context. All the other information will be assigned to the entities through objectified relationships that are defined through our extended version of the *IfcRelationship*-class, which makes the information distinguishable.

5. Conclusions

The current release 2x of the IFC specifications have limitations in modeling as-designed and as-built information into one project model to support the automation of as-built conditions. It is necessary to extend the current IFC representation without increasing the complexity unnecessarily from both an understandability and a processability point of view.

For the purpose of allowing IFCs to represent both design and as-built information in one project model simultaneously, we propose a solution that allows for fast processing of the IFC model without increasing the complexity of the IFCs. This was accomplished through the utilization of the *IfcRepresentationContext* concept in IFCs and adding a new attribute to the *IfcRelationship* class.

We have tested this approach in automating the as-built assessment of the construction of a steel frame that we observed on a recent project. We will further test the scalability of this concept by adding different types of design, schedule and as-built information that we are going to collect in an upcoming construction project.

6. Acknowledgements

The project is funded by a grant from the National Science Foundation, CMS #0121549. NSF's support is gratefully acknowledged. Any opinions, findings, conclusions or recommendations presented in this paper are those of authors and do not necessarily reflect the views of the National Science Foundation.

7. References

[1] Cyra Technologies, Inc. *Case Studies*; http://www.cyra.com/case_studies/case_studies.html (accessed: 06/2002)

[2] Cyra Technologies, Inc. *Cyra Case Studies: Condition Assessment of Old Suspension Pipeline Bridge*; 1999; http://www.cyra.com/PDFs/PipelineBridge.pdf (downloaded: 06/2002)

[3] The Steel Construction Inst. *CIMsteel Integration Standards, Rel. 2, Vol. 1, Overview*;http://www.cis2.org/documents/Vol1/SCI-P265on.pdf (06/2002)

[4] R. J. Hitchcock, M. A. Piette, S. E. Selkowitz: *Documenting Performance Metrics in a Building Life-cycle Information System*; Building Technologies Depart., Lawrence Berkeley National Lab., Univ. of California, Berkeley, CA, August

1998;http://eande.lbl.gov/BTP/papers/41942.pdf (downloaded: 06/2002)

[5] IAI: *IFC 2x Extension Modeling Guide*; http://cig.bre.co.uk/iai_uk/documentation/Ifc2x_EMG/EMG-Base.htm (06/2002)

[6] T. Froese, M. Fischer, F. Goebler, J. Ritzenthaler, K. Yu, S. Sutherland, S. Staub, B. Akinci, R. Akbas, B. Koo, A. Barron, J. Kunz: *Industry Foundation Classes for Project Management - A Trial Implementation;* ITcon Vol. 4 (1999); http://www.itcon.org/1999/2/paper.pdf (downloaded: 06/2002)

FIGURES:

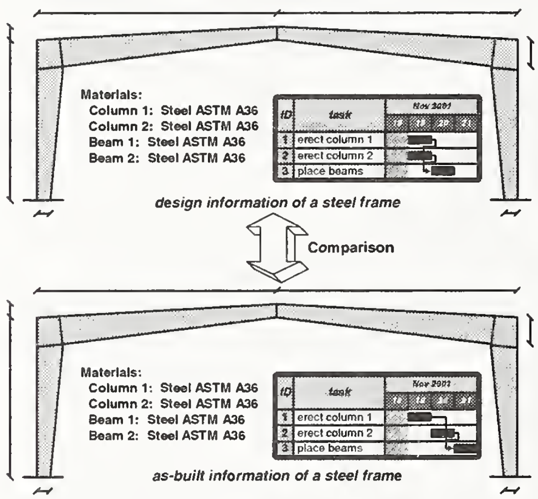


Figure 1: The design and as-built information of a steel frame.

Figure 2: Representation of material information of a column.

Figure 3: Representation of element type information through property sets.

Figure 4: Representation of element type through predefined types.

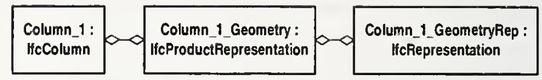


Figure 5: Representation of geometric information as a group of representation items.

Figure 6: Representation of geometric information through property sets.

Figure 7: Representation of scheduling information to a column.

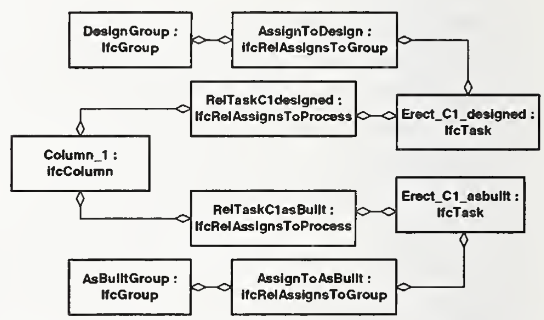


Figure 8: Representation of design and as-built information through groups.

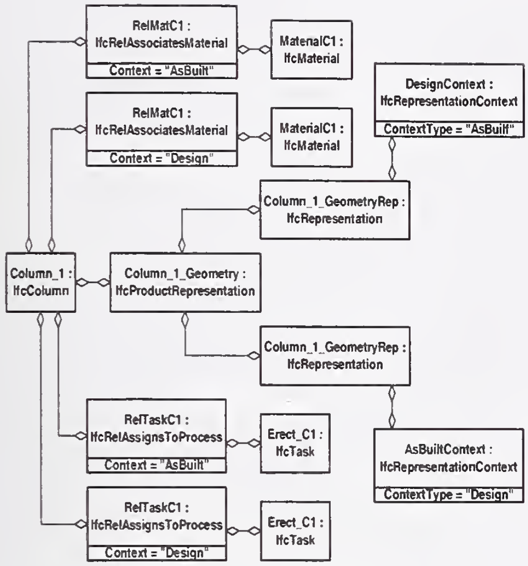


Figure 9: The proposed concept.

A Case Study: Using The Wearable Computer In The Construction Industry

by

Scott Fuller, Zhihui Ding, and Anoop Sattineni

ABSTRACT: This paper aims to explore the prospective applications of cutting-edge technology like wearable computers in construction industry. The history, current research, application areas and future trends of wearable computing are investigated. The first-hand experience of the authors through a well-designed field experiment is introduced in detail to show how the technologies actually work. The field experiment is performed using three different methods, including by hand, by Palm PC and by wearable computer. The pros and cons for each are analyzed and compared to reveal why wearable computer is a better tool to help construction professionals to perform their daily duties like "punch-list". The results in summary showed that in immediate future, wearable computer is a better solution for construction than Palm PCs. But the advocates of the wearable computer should always bear in mind that the technologies are always developing. Palm PCs are still expanding their functionality while maintaining the mobility. It is not impossible that tomorrow they could have the same storage and processing capacities as wearable computers today. So Palm PCs are going to be strong competitors of wearable computer.

KEYWORDS: Construction, Palm PC, Punch List, Wearable Computers

1. INTRODUCTION:

The construction industry is on the edge of something that could drastically improve the entire construction process. Recently, a new kind of computer has caught the public attention, a 'wearable computer': a computer that can be worn on a human body. Its mobility and functionality has made it useful in many places where a regular computer could not be carried. This "...has the ability to bridge time/distance constraints imposed by working at remote construction projects...the opportunity for real-time decision making in construction is improved." [1]

2. LITERATURE REVIEW

Through the extensive research, the authors found the term "wearable computer" has been defined in several ways. For example, according to MIT wearable computing lab, wearable computer is a computer that should be worn much as "eyeglasses or clothing", and interact with the user based on the context of the situation. With features like heads-up displays, unobtrusive input devices, and personal wireless local area networks, the wearable computer can act as an intelligent assistant. NASA (National Aeronautics and Space Administration) used the term Body Wearable Computer (BWC), which is a battery-powered computer system worn on the user's body. And the unit is designed for mobile and predominantly hands-free operations, often incorporating head-mounted displays and speech input.[2]

Today, there are a lot of commercial wearable computing systems available in the market. Some prototypes require special glasses, while others rely on sensors similar to medical electrodes on various body parts to feed data to the computer. The differences depend on the purpose of each system. Among all systems, Xybernaut Corporation is one of the largest sellers, partly because it has locked up a big

[1] Mills and Beliveau, 1998

Auburn University, Department of Building Science, 119 Dudley Hall, Auburn, Alabama 36849, 334-844-5431, fulleds@auburn.edu

[2] Mizrachi, 2000

chunk of the patents.[3] Figure 1 shows Mobile Assistant V, one member of Xybernaut wearable computer family.

In summary, wearable computers are more powerful, durable than laptop/palm computers, but more convenient to use than desktop PCs.[4] As Ed Newman, chairman and CEO of Xybernaut, said, “We're not talking about a Windows CE device; we're not talking about a PDA, we're talking about a product that literally will run anything your laptop or desktop will run with almost no change." Wearable computers are able to run “all the programs on a PC”, including word processing, spreadsheets, and databases, but in a more efficient way.[5]

Today, wearable computers are not stagnating at the laboratory and exotic-toy stage.[6] Instead, they are finding wide use in the real world. For Xybernaut alone, its current customers include many companies that have large field forces, such as FedEx Corp., Bell Canada, and the U.S. Army and Navy, to name a few.[7]

FIATECH, a research and development firm, has a two-year goal to introduce wearable computers to the construction industry. FIATECH will record the activities of people in the field to understand how their work is done and how it can best be improved through the use of a wearable computer.[8]

One of the major issues with the construction industry is that the production activity is dispersed and the site locations frequently change. This is a disadvantage for the construction industry that most other industries do not have. This makes IT support and integration for the construction industry more difficult.[9] This in turn makes it a great candidate for the wearable computer or mobile computing. In addition, effective communications is the most important component for successful projects. Communication between the participants on a job site is difficult, due to the large number of participants, their remote locations, the visual nature of many issues/solutions, the number of problems that arise at the site and the need for them to be solved on-site. However, it is also exactly these kinds of issues that might make the wearable computer a very effective, efficient, useful and beneficial tool for construction.[10]

Certainly for the construction industry the wearable computers and all attached devices must be accessible, lightweight, and easy to use. They must be able to withstand all types of weather and environmental conditions. For example, the screen should be readable in all types of lighting, the computer should be impact resistant, have rechargeable batteries and a host of other functions and ergonomic features. [11]

[12]Mills and Beliveau proposed the idea of “Virtual Site Visit” using wearable computers. A "Virtual Visit" is a visit to a construction site to observe, evaluate, clarify and correct actions and activities, but it occurs without the individual's actual physical presence. This is of great benefit when it is impossible to visit the site and will diminish project delays. The reduction in travel time and increased project knowledge translates into immediate cost savings to the contractor, owner and designer.

Just as Jim Porter, with Dupont says, “In today’s competitive environment, we need to be able to access data, drawings, and other information on demand remotely at the construction site. Wearable computers will allow us to do this.” [13]

Essential to using advanced computing technology for field decision making is an understanding of the jobsite informational needs. Research on using wireless communications for construction information needs has been undertaken and reported by the Construction Industry Institute.[14]

Two of the authors conducted a preliminary survey with some construction firms. The

[3] Maney, 2001
[4] Merritt, 2001
[5] Mizrachi, 2000
[6] Mizrachi, 2000
[7] Neel, 2000
[8] Newland and Owings, 2001
[9] Rebolj, Magdic, and Cus-Babic, 2000

[10] Miah, et al, 1998
[11] Newland and Owings, 2001
[12] Mills and Beliveau, 1998
[13] Newland and Owings, 2001
[14] De la Garza, et al, 1997

purpose of the survey was to investigate the potential uses for the wearable computer technology and the interest in the industry.

The survey indicated that the construction industry is already using information technology to improve performance on projects. It also showed that an enormous amount of data is recorded on the job site by construction personnel. All respondents indicated that they would be willing to experiment with wearable computer technology, thus indicating that there is a tremendous potential for the success of wearable computer technology.

3. METHODOLOGY

The objective of this research project is to test the wearable computer in a real-life situation in construction to see how technology could help construction professionals. For example, compare the different time periods required to perform the same amount of work, under different methods including by hand, by Palm PC or by wearable computer.

In order to do so, first the author should have a wearable computer system configured and use it extensively to get to know it well. Secondly, the author should identify a testing area where experiments with such appropriate scale could be performed to "mock" the real-life situation. Lastly, putting the two aspects together, the author should design a testing scheme and further configure the wearable computer system for the testing purpose.

The wearable computer used in research will come from Department of Building Science, Auburn University. This Xynernaut computer, the Mobile Assistant V (MA V), will contain a 500-MHz Intel Celeron processor, 128-MB of RAM, and a 2-GB hard drive that can be worn in a vest or on a belt. It will be able to run all major PC operating systems, including Windows 2000/NT and Linux. The MA V will also incorporate a daylight-viewable, head-mounted display or a wrist-mounted SVGA flat panel touch-screen viewable in all light levels. Interaction with the computer will be made via a wrist-worn keyboard, touch-screen, or voice recognition software.

The following lists three ways that we proposed to test the wearable computer for construction applications: 1. Traditional Manual Method, 2. Palm-size PC, 3. Wearable computer.

In our testing, the actual situation might vary from the above. However, at least, the following conditions must be met, before performing research related to punch list: a building should be identified and available for research purpose; drawings (traditional blueprints or electronic format), specifications (or scope statement) should be made available; the author should develop sample checklist and punch list for field use; technical tools, such as Palm-Size PC, wearable computers, should be well-configured before hand; and bring some accessories, like stop watch to record time, pens and paper.

4. RESULTS

4.1 Field Testing

The author selected Intramural Fieldhouse at Auburn University as the target building for carrying out the "Punch-list" experiment. Based on the information from drawings and specifications, a sample checklist and punch-list were first developed in MS Excel spreadsheet and Word document format, respectively. Then three trips were made separately to perform the building inspection, by hand, by Palm PC and by wearable computer. Plumbing and electrical systems were the primary emphases, but interior finishes and furniture were also included. The process is defined as from the time when the author entered the building to the time when the checklist and punch list were completed.

From the experience in performing the "punch-list", the author summarized the pros and cons for each method, as in Table 1.

Firstly, the advantage for using mobile computing device over manual method is quite apparent. Whether using Palm PC or wearable computer, when the punch list is done, it will be sent through e-mail instantly. But with the manual method, the superintendent will have to come back, find a desktop PC, and manually put all information into a computerized system (or scan in the paper form) to be sent via e-mail to subcontractors at the same day.

If we assume the distance between the jobsite and home office is two miles, the additional travel time and data input time, in our case, will be approximately 20 minutes. In real life, the travel time will be even significantly longer. So if assuming the added time for communicating the punch list to subcontractors is one hour under the manual method, the above timetable should be adjusted as Table 2 indicates. Now one could see clearly that using computerized methods, the problem of "double data entry" and delivery time could be eliminated, thus significantly improving the efficiency.

In our experiment, the electronic drawings are around 20 MB in total. This was not a big issue for our wearable computer, which has two hard drives (2 GB in total storage capacity). But it was a huge burden for our Palm PC, which only has 8 MB storage. So our experiment reveals that wearable computer did outperform Palm PCs in storage capacity, with or without external devices.

4.2 Data Access

With wearable computer, the electronic drawings were loaded as well as electronic version of checklist and punch-list, all that the author needed to do is to carry the wearable computer around. It should be noted that in our experiment, the idea of electronic drawings didn't apply to the manual method or Palm PC at all. That is because Palm PCs do not have the power to manipulate electronic drawings. Even within the near future, this situation is not very likely to change, considering the fact that the drawing viewing tool itself alone will take more than 4 MB and doesn't work with Palm PC at all.

4.3 Data Processing

Wearable computers don't require special applications loaded in order to process data. But Palm PCs will require special software to perform the same tasks., like Pocket Microsoft Office (Pocket Word, Pocket Excel, Pocket Access etc). This software costs the same price, if not more, as the standard software does. In addition, some features in regular applications will be unavailable in the software designed for Palm PCs. In this experiment, the authors was able to add a password-protected digital signature to the punch list when using wearable computer. But using Palm PC, the document was converted to Word Pad text file, and all macros were lost. In addition, the software installation process is also more complicated for Palm

4.4 Data Synchronization

An important issue for mobile computing devices is how the data on them could be communicated back and forth into the current corporate network, which is almost exclusively based on a group of desktop workstations. Although our experiment finished when punch lists were complete, in real life, all data on wearable computer or Palm PC must be finally backed up and updated. E-mail is fine for exchanging small files like the punch list as a Microsoft Word document in our experiment. But for larger files, other solutions must be sought.

For Palm PC, as described earlier, this relies on the communication between Palm PC and a desktop PC. First prepare desktop PC for synchronization purpose by installing special utility software like Microsoft ActiveSync on it. Then connect Palm PC with desktop PC by using a cable. Activate Microsoft ActiveSync or other similar program, and the data between the two will be exchanged or updated. Except for the first time, which requires software installation, all one has to do is connect the cable and click the button.

In our experiment, the solution for wearable computer was "Sneak Net", i.e. transporting data via 3-1/2" floppy-disks with the help of related USB devices. But even by that, wearable computers still outperform Palms in this aspect, because the USB devices are basically mechanical operated (Plug-Play-Unplug) and easy to bring with wearable computers.

From the discussion above, it seems that within the near future, wearable computer is a better solution for construction than Palm PCs. But the advocates of wearable computer should always bear in mind that the technologies are always developing. Palm PCs are still expanding their functionality while maintaining the mobility. It is not impossible that tomorrow they could have the same storage and processing capacities as wearable computers today. So Palm PCs are going to be strong competitors of wearable computer.

5. CONCLUSIONS:

For wearable computers, a lot of things could be done. Especially, one should look at the strengths of Palm PCs to improve the performance. For example, wearable computers could be made smaller in size, easier to carry, having lighter CPU or even integrating CPU with monitor into one unit etc.

In our experiment, the punch list and checklist files were first edited in a desktop computer, and then saved to a floppy disk to be later transferred to wearable computer. This is the "Sneak Net" process we talked about earlier. But in construction, the needs for data updates between office and field could be very frequent and heavy. Thus the data synchronization process described above could be so labor intensive and time consuming for anyone.

Ideally, the wearable computer will be a workstation ("Client") in the corporate client-server networking environment. Just like a normal desktop PC, a wearable computer could download file from, and also upload files to, the company server. It will have Internet access via the corporate server. Considering the diverse nature of the construction industry, the wearable computers could be configured into wireless communication mode. The limitations for current wireless technology could be overcome by setting up a small-scale LAN in each jobsite, where the local server in the trailer could directly meet the needs of mobile devices by communicating the company server. However, the costs for establishing such networks and maintaining them could be so prohibitively expensive for any contractor. Some recent innovations in IT industry like ASP (Application Service Provider) might provide better solutions to such problem.

In addition to the issues discussed above, some technologies that are supposed to work with wearable computers, like speech recognition and head-mount display, are far from mature to be applied in construction industry in a significant scale.

In summary, the application of wearable computers in construction industry could offer improved efficiency. But the future researchers and developers should improve the performance and practicability of the technology before it is widely accepted by construction industry professionals.

6. REFERENCES:

De la Garza, Jesus M., Howitt, Ivan. *Wireless Communication and Computing at the Construction Jobsite*. Construction Industry Institute Research Summary 136-1, 1997.

Maney, K. (2001, April 18). *Wearable PCs: A different kind of fashion statement* [WWW document]. URL http://www.usatoday.com/life/cyber/tech/review/2001-04-18-wearable.htm

Merritt, R. (2001). Switching fabric. *Electronic Engineering Times, 1191, 1-3*

Miah, T., Carter, C., Thorpe, A., Baldwing, A., and Ashby, S. (1998). Wearable computers—an application of BT's mobile video system for the construction industry. *BT Technol Journal*, Vol. 16, No. 1, pp 191-199.

Mills, T., and Beliveau, Y. (1998). *Virtual Site Visit Using Wearable Computers*. The 1998 International Conference on Wearable Computing, Fairfax, VA, 1998.

Mizrachi, D (2000) Info-Fashions for the New Millennium, *Online*, 24(1), 82-85

Neel, D (2000) Wearable PC goes to work, *InfoWorld*, 22(47), 10

Newland, B., and Owings, C. (2001, Jan. 9). *In the field, wearable computers will change how work is done*. [WWW document]. URL *http://www.djc.com/news/co/11127075.html*

Rebolj, B., Magdic, A., and Cus-Babic, N. (2000, Jan 9). "Mobile Computing in Construction". [WWW document]. URL *http://kamen.uni-mb.si/cgi/Mobile%20computing%20in%20construction.htm*

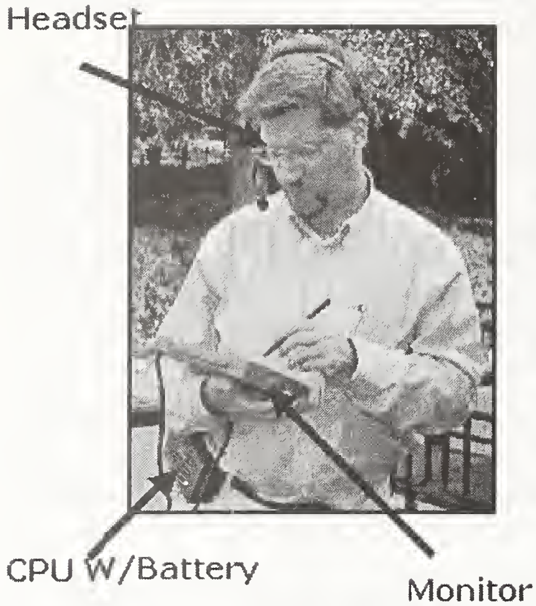


Figure 1: MA V Wearable Computer by Xybernaut Corporation.

Table 1.
Comparison of Three Different Methods

Method	Advantages	Disadvantages	Human Factor
Manual	Quick and easy to input data	Hard to manipulate blueprints and writing pad simultaneously	Fatigue in wrist and arms and sweating body
Palm PC	No need for paper forms, Palms easy to carry, light weight	Still require blueprints, screen hard to read where not much lights, slower input	Eye fatigue
Wearable Computer	No need for blueprints or paper forms, instant access to everything from anywhere	Require time to put up parts on body, and wait for the system to boot up; Require operations to drawings like zooming in and out, moving up and down	Finger fatigue after continuous data input using arm–worn keyboard

Table 2.
Adjusted Results due to communication time.

Method	Original Time	Adjusted Time
Manual	76 Minutes	136 Minutes
Palm PC	95 Minutes	95 Minutes
Wearable	66 Minutes	66 Minutes

The Value of Handheld Computers in Construction

by

Kamel S. Saidi[1], Carl T. Haas[1], Nicole A. Balli[1]

ABSTRACT: Construction is an information intensive industry in which the accuracy and timeliness of information is paramount. Construction projects can experience extensive delays or rework due to information that is unavailable, inaccurate or simply outdated. Handheld computers (HHC) have the potential to solve some of these problems by providing field workers with accurate, reliable and timely information at the location where it is needed. Thus, HHC's can increase the amount of direct work on a project indirectly by directly decreasing the time spent on support work (such as accessing drawings and sending RFI's) and by reducing idle time. Applying a HHC evaluation method to 6 hypothetical construction field activities (punchlisting, materials tracking, MSDS access, drawing access, RFI's, and quantity surveying) showed that HHC's could potentially save time and improve accuracy at the task and activity levels of a construction project. However, barriers related to the HHC's technological limitations and to the nature of the construction industry must be overcome in order to reap the full benefits of HHC's.

KEYWORDS: field data, handheld computers, material tracking, punchlisting, quantity tracking

1. INTRODUCTION

1.1 Problem Statement

The successful and timely completion of a construction project depends on the accuracy and timeliness of a vast amount of information [1, 2]. Craft foremen spend more than 50% of their time in the field where data is difficult to access outside of the site office. Projects often experience extensive delays or rework due to information that is unavailable, inaccurate or outdated. These delays decrease the overall productivity of the project and increase indirect costs due to schedule delays or direct costs due to rework. The construction industry is in need of tools that can provide accurate, reliable, and timely project information *to* the field and gather and transmit up-to-date project information *from* the field. Handheld computers (HHC) can potentially fulfill these needs.

1.2 Research Objective

The objective of this research was to investigate the potential of HHC's to add value to a construction project through impacts on time and money[2] and to evaluate this potential.

1.3 Hypothesis

HHC's can *indirectly* increase direct work by *directly* decreasing the amount of support work and idle time within an activity (see Figure 1).

1.4 Methodology

The following methods contributed to this research: 1) an extensive literature review was performed; 2) informal interviews with construction contractors and IT companies were administered; 3) a simple, systematic HHC evaluation method was developed; 4) HHC hardware, software and other peripherals with specifications suitable and beneficial to the construction industry were classified; 5) six construction activities in which HHC's were thought to have the greatest potential benefit were identified; and 6) an evaluation method was applied to the above activities as case studies. Details of the above methods and their results are presented in [3, 4].

2. BACKGROUND

A HHC is a self-contained electronic device that fits in the palm of a user's hand and possesses, at a minimum, enough computer processing power to surpass the functions of an electronic personal organizer and to run software applications that can extend its built-in functionality.

---

[1] Department of Civil Engineering, University of Texas at Austin, ECJ 5.200, Austin, TX 78712, (512) 471-4645, kamel@mail.utexas.edu, haas@mail.utexas.edu, nballi@mail.utexas.edu

[2] Other measures such as safety, rework, and productivity are also widely used, however, these are intrinsically related to the above three metrics. Although quality is as important as the other two metrics, apart from measuring factors such as rework and owner satisfaction in an effort to quantify it, quality remains an illusive metric.

The use of HHCs on the construction jobsite was investigated as early as 1992 for field data acquisition [5]. The implementations of HHCs in construction discussed in the literature have focused primarily on project management, schedule management, facility inspection, and field reporting applications [4]. Various construction firms have started using handheld computers on the jobsite for gathering schedule, quality, layout, inspection, and other types of information [6, 7, 8, 9]. However, due to the relative immaturity of HHC use in construction, there have been very few applications in construction that may be considered an accepted way of doing business.

A study of the software applications available on the handheld computer market conducted in 2001 showed that there were approximately 40 titles geared specifically toward the construction industry whereas over 300 titles were commercially available for the health industry alone [4]. This seems to indicate a lack of interest in HHCs on the construction industry's part [10] and a lack of interest in developing applications for the construction industry on the part of the HHC hardware and software manufacturers. Of the top 8 HHC manufacturers contacted by the authors, none indicated that they had identified the construction industry as a differentiated customer for their product development and marketing [4]. In contrast, the manufacturing, white goods, process plants, transportation, healthcare, and other industries have been marked as targets by most of the same manufacturers.

In a survey of 179 construction foremen, Alemany [11] showed that foremen who used computers at work saved time on paperwork and spent more time on supervision. Most of the surveyed foremen expressed a desire to automate time reporting, visualizing and interpreting drawings, job progress recording, and tools and materials management functions [11].

In another survey conducted internally by a large construction company the authors found that supervisors spent between 36 to 50% of their time on paperwork related to employee time keeping and material management functions [5]. The above 2 surveys suggest that using HHCs effectively in the field for employee time keeping and materials management alone could enable foremen to spend more of their time supervising. Consequently, this could have positive impacts on productivity and quality. Similarly, providing construction workers with HHCs that can help them locate tools, equipment, and materials, send requests for information (RFI's), and access relevant schedule information (among other important functions) could potentially allow them to spend more time on direct work and less idle time waiting for answers or needed tools and materials. Other benefits of HHC in construction have also been identified in the literature [6, 12, 13, 14].

3. THE EVALUATION METHOD

Several researchers have proposed formal techniques for evaluating IT in construction [15, 16, 17, 18]; however, none of these techniques deal with HHCs. The justification for using HHCs in the construction industry (and other industries) must account for impacts on the organization's IT infrastructure, the construction processes, and so on [14, 19].

3.1 Basis of the Method

Since most technologies are applied at the task level within a project [20], and their impacts propagate up toward the project level, the evaluation of the suitability of using HHCs on a project must begin at the task level. The HHC evaluation method presented herein breaks down a construction activity hierarchically into a detailed set of final elementary tasks [21], and defines time and cost values for each elementary task [22]. As a means of representing the decomposed task hierarchies, information flow charts (also known as decision-action diagrams, logic diagrams, process flow charts, etc.) are also used [21, 22]. Finally, the evaluation method incorporates a simple accounting process whereby elementary task times and costs are accumulated in order to calculate totals [9].

3.2 The Evaluation Process

The evaluation method first requires that the construction process as it currently exists (i.e., the traditional process) be systematically documented. While the traditional process for the same activity may differ from company to company, the change that would occur in the process when HHCs are introduced is the evaluation method's primary concern.

Once the elementary tasks of the construction activity in question are defined, the next step involves assigning responsibilities for each task. From the list of elementary tasks, a flow chart is created in order to capture the sequence of activities and any feedback loops that may exist. Next, minimum and maximum completion times are assigned to each elementary task to capture the variation that may occur. Although, the times assigned may not be entirely accurate, or may differ between companies and/or projects, as already stated, the differences in task times between the traditional process and the HHC process are the focus. Finally, the possible errors and corresponding delay times for each elementary task are documented.

After describing the traditional activity, the same method is applied to the activity as it would exist with the introduction of HHCs. Depending on the activity, certain tasks are changed, combined, or altogether eliminated. Times are reassigned to each elementary task in the HHC process and the potential errors and associated delays are also adjusted.

The final step in the evaluation process is to estimate the total activity time both with and without the use of a HHC. The time difference between the traditional and HHC processes is estimated to determine whether the use of the HHC might be beneficial to that particular activity. Table 1 shows a blank sample form that is used to record task information for each activity.

4. CASE STUDIES

The HHC evaluation method outlined above was applied to 6 construction field activities: 1) punchlisting, 2) materials tracking, 3) MSDS access, 4) drawing access, 5) RFI's, and 6) quantity tracking. Except for the quantity tracking activity's evaluation (which was based on field observations and interviews) the method was applied to theoretical models of the construction activities involved rather than to actual activities on a construction project. The reader is referred to [3, 4] for detailed results and descriptions of these evaluations. A primary assumption made in the case studies below is that the introduction of HHCs in each activity is coupled with an implementation of the Center for Construction Industry Studies' (CCIS) Tier II strategy [22]. One of the central premises of the Tier II strategy is that field personnel have greater access to information and certain decision-making powers without management's approval.

4.1 Punchlisting

The punchlisting activity lends itself well to HHC implementation because it is a field-based activity whose information is typically collected into a form. In addition, the punchlisting process is cyclical, since it may occur repeatedly throughout the project and some items may be re-listed on the punchlist if not satisfactorily completed [23].

Applying the HHC evaluation method to the punchlisting activity showed that the use of HHCs can theoretically eliminate 14 elementary tasks, which could reduce each punchlisting cycle's time by an estimated 40%. The addition of HHCs to the punchlisting process can also contribute to a 39 to 46% reduction in delay time. Overall, HHCs can potentially reduce delay time by approximately 50 to 70%.

4.2 Materials Tracking

Received materials are often improperly recorded, relocated or not recorded at all. Materials that are lost, misplaced or improperly stored can cause major delays and disruptions on a project and drastically affect project cost and schedule [24]. Handheld computers could help resolve some of these problems by eliminating handwritten notes and the reliance on human memory, and by offering foremen access to up-to-date material information. The materials tracking activity was selected because it is field-based and was identified by foremen as a priority for automation in Alemany's [11] research.

Applying the HHC evaluation method to the materials tracking activity showed that the use of HHCs can potentially eliminate 9 elementary tasks. Approximately 26 to 51% of the overall activity time can also be saved by implementing HHCs and the Tier II strategy in concert.

Overall, the potential delay time saved was estimated to be 88 to 95%, with a majority stemming from the addition of a HHC to the process.

4.3 MSDS Access

The MSDS (Materials Safety Data Sheets) access activity was selected because it requires onsite access to large amounts of textual information, and thus lends itself well to HHC implementation. For this evaluation it was assumed that an online MSDS database could be accessed wirelessly (many are currently publicly available) or stored on the HHC itself.

Applying the HHC evaluation method to this activity showed that 5 elementary tasks can be theoretically eliminated. Implementation of the Tier II strategy eliminated those tasks that required information transfer between different hierarchical levels of the organization; while the addition of HHCs to the process eliminated travel and distribution tasks. A 59 to 71% reduction in overall activity time was estimated. The total reduction in the activity delay time was approximately 65 to 75%.

4.4 Requests for Information (RFI)

The RFI activity was selected for investigation because a large number of inquiries arise at the work face and a means of documenting them and receiving answers quickly can eliminate delays. In addition, the process does not involve much data entry and can take advantage of on-site wireless communications to send and receive information. The new HHC process assumes that the Tier II strategy would allow the foreman to communicate directly with the A/E via e-mail, and that the work in question could be viewed in a digital photograph (sent via e-mail wirelessly from the HHC) by the A/E rather than in person.

Applying the HHC evaluation method to the RFI activity showed that the use of HHCs can theoretically eliminate only one elementary task. However, the new process reduces the activity time by an estimated 16 to 23%. While all of that time is saved due to the use of a HHC, the time saved due to implementation of the Tier II strategy is captured in the delays. The reduction in delay time is key in this activity and is due in large part to the elimination of the hierarchical structure that a traditional RFI follows. In addition, the delay between the time that an RFI is generated and answered is greatly reduced because the architect does not have to travel to the site. The new process can potentially reduce delay time by 83 to 91%.

4.5 Drawing Access

The drawing access activity was chosen because it is a field-based activity that requires only access to information, and therefore does not require any data entry. It is assumed that foremen will have access to a central database to which drawings are regularly uploaded. In addition, the assumption is made that in the new process the Tier II strategy will allow the foreman access to the drawing database directly rather than having to go through a superior.

Applying the HHC evaluation method to the drawing access activity showed that the use of HHCs could potentially eliminate 3 elementary tasks. Using the HHC and the Tier II strategy in this process reduces the activity time by an estimated 70%, primarily due to reductions in travel time and time taken to obtain the information through hierarchical channels. Delays associated with the eliminated tasks are also reduced by an estimated 64 to 72% in this activity. While the overall result shows a reduction in delay time, the issue of the HHCs small screen could prove to have adverse effects on the activity's productivity.

4.6 Quantity Tracking

The purpose of tracking (or surveying) quantities on a construction project is to measure progress in order to control cost and schedule. The quantity surveyor (or tracker) tracks the quantities of materials that have been installed or that are in various stages of installation. These quantities are essential for controlling a project's cost, schedule, and quality. Therefore, the accuracy of the quantity survey data is critical [4].

Applying the HHC evaluation method to the quantity tracking activity showed that the use of HHCs eliminated 6 elementary tasks and saved approximately 60% of the overall activity's time; in addition to improving the accuracy of the data, providing an auditing tool to check takeoff quantities in the field, and reducing the chance of quantity over-reporting.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The use of HHCs in six construction activities was evaluated above. The evaluations showed that time savings at the task and activity levels do not translate directly into project-level time savings and that benefits are more likely to be achieved if HHCs are implemented in multiple activities and projects.

Based on the preliminary experiments with a HHC purchased as part of this research (see [4]), HHCs are currently bound by several key technologies that limit their functionality under certain conditions. These limitations involve HHC features such as screen size, screen visibility, processing capability, and input method. Table 2 presents a list of construction tasks that *are* suited for HHCs, followed by tasks that are *not* suited (these tasks do not take into account HHC's extended range of functions when combined with other peripherals).

This research also found that the barriers to HHC implementation in construction are a result of two factors: 1) the HHC technology's limitations and 2) the construction industry's characteristics. The HHC technology's limitations where discussed. The construction industry barriers consist of the physical jobsite conditions (such as temperature, humidity, dust, etc.) as well as organizational issues such as the industry's fragmentation and low risk tolerance, among others [4].

Handheld computers have many benefits that can improve construction processes. The most significant benefit is perhaps the HHC's ability to provide workers with real-time access to relevant information at the jobsite, and to send real-time information back from the jobsite to the appropriate decision makers. In addition, an HHC's ability to improve the accuracy of the information being exchanged is one of its primary added values in construction. The type of information and the transmission method are some of the issues that must be assessed during the design of an HHC evaluation and implementation strategy.

5.2 Recommendations

The lack of empirical data on HHC performance in construction could be improved through well-documented pilot projects at construction companies and through controlled experimentation with HHCs under simulated environments. In addition, future research should also address HHC hardware issues that constitute barriers to their implementation on construction projects.

6. REFERENCES

1. Rezgui, Y., Cooper, G., and Brandon, P. (1998), "Information Management in a Collaborative Multifactor Environment: The COMMIT Approach," *Journal of Computing in Civil Engineering*, 12 (3), 136-144
2. Shahid, S. and Froese T. (1998), "Project Management Information Control Systems," *Canadian Journal of Civil Engineering*, v 25, n 4, 1998, pp 735-754
3. Balli, N.A. (2002), "Handheld Computer Applications for the Tier-II Worker," Masters Thesis, The University of Texas at Austin, Austin, Texas
4. Saidi, K. (2002), *Possible Applications of Handheld Computers to Quantity Surveying*, PhD Dissertation, The University of Texas at Austin, Austin, TX
5. McCullouch, B.G. and Gunn, P. (1993), "Construction Field Data Acquisition with Pen-Based Computers," *Journal of Construction Engineering and Management*, v119, n2, pp374-384
6. Inglesby, T. (2000) "Handheld Computing: In the Field, Handhelds Can Make a Difference," *ConstrucTECH*, 3(4), http://www.constructech.com/, accessed 7/31/2000
7. CII (2000) "Mobile Inspection Assistance," CII's Emerging Construction Technologies website, http://www.new-technologies.org/, accessed 8/1/2000
8. CII (2000) "Palm Estimating," CII's Emerging Construction Technologies website, http://www.new-technologies.org/, accessed 8/1/2000
9. CII (2000) "Updater: Mobile Schedule Tracking Software," CII's Emerging Construction Technologies website, http://www.new-technologies.org/, accessed 8/1/2000
10. Coble, R.J. (1995), "Expanding Pen Computing Capabilities to Include More Effective Field Application," *Computing in Civil Engineering*: proceedings of the Second Congress held in conjunction with A/E/C Systems '95 / ASCE, Atlanta, GA, pp1340
11. Alemany, C.H. (1999), "Construction foremen computer use and impacts," Masters Thesis, The University of Texas at Austin, Austin, Texas
12. Williams, T.P. (1994), "Applying Portable Computing and Hypermedia to Construction," *Journal of Management in Engineering*: May-June 1994, v10, n3, pp41
13. Newell, D.M. (1994), "Pen Computers in Civil Engineering," *Computing in Civil Engineering*: proceedings of the First Congress held in conjunction with A/E/C Systems '94 / ASCE, Washington, D.C., June 20-22, pp1631
14. de la Garza, J.M. and Howitt, I. (1998), "Wireless Communication and Computing at the Construction Jobsite," *Automation in Construction*, v7, pp 327-347
15. Lindorfs, C., Chang, P., and Stone, W. (1999), "Survey of Construction Metrology Options for AEC Industry," *Journal of Aerospace Engineering*, v12, n2, pp 58-64
16. Chun, D.A., Seo, J., and Haas, C.T. (1997), "Leveling Technologies for Building Construction," Proceedings of the ASCE Construction Congress V, pp1030-1037, Minneapolis, MN., October 5-7, 1997
17. Jung, Y. (1997), Information Systems Planning Methodology for the Construction Industry, Dissertation, The University of Texas at Austin, Austin, TX

18. Jung, Y. and Gibson, G.E. (1999), "Planning for Computer Integrated Construction," *Journal of Computing in Civil Engineering*, ASCE, v13, n4, pp217-225
19. Aberdeen Group, Inc. (2000), "The Benefits of Mobilizing Enterprise Applications with Handheld Devices," Aberdeen Group, Inc., Boston, MA, August 2000
20. Back, W.E., Moreau, K.A., and Toon, J. (1998), "Determining the Impact of Information Management on Project Schedule and Cost," *CII Research Report 125-11*, Construction Industry Institute, The University of Texas at Austin, Austin, Texas
21. Kirwan, B. and Ainsworth, L.K. (1992), *A Guide to Task Analysis*, Taylor & Francis Ltd., Washington, DC
22. Borcherding, J.D., Glover, R.W., Haas, C.T., and Tucker, R.L. (2001), "Metric-Based Implementation of Tier II Work Force Strategy," Center for Construction Industry Studies, Report No. 20, The University of Texas at Austin, Austin, Texas, August 2001
23. Oberlender, G.D. (2000), Project Management for Engineering and Construction, The McGraw-Hill Companies, Inc., Boston, MA
24. Stukhart, G. (1995), *Construction Materials Management*, Marcel Dekker, Inc., New York, NY

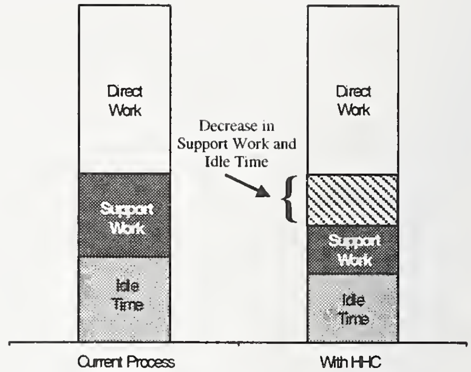


Figure 1. HHC's decrease support work and idle time.

Table 1. An activity task information form

Task ID	Task Description	Labor	Task Time (minutes)		Possible Delay (minutes)		Source of Delay
			min	max	min	max	
10							
20							
30							
40							
...							

Table 2. Tasks for which HHC's *are* and are *not* suited.

#	Tasks that *are* Suited	Example
1	Tasks that require access to large amounts of text information	Reading MSDS sheets, building codes, knowledge base, etc.
2	Tasks that require viewing a small detail of a document	Viewing a close-up of a steel beam connection diagram
3	Tasks that require the entry of binary data	Answering yes/no questions, checking-off items on punch lists
4	Tasks that require the entry of data into a form	Filling-in a safety or equipment usage report, recording material receiving information, etc.
5	Tasks that require instant transfer of small amounts of information to and from a network	Sending and receiving e-mails, looking up the latest material procurement information
#	**Tasks that are *not* Suited**	**Example**
1	Tasks that require computer processing power comparable to that found in desktop computers	Editing a 3-D construction drawings
2	Tasks that require a "big-picture" view of a document	Viewing a drawing or a network schedule
3	Tasks that require a constant (i.e., always on) connection to a computer network	Working with data stored on a mainframe
4	Tasks that require a considerable amount of manual data entry (or writing)	Writing a progress report
5	Tasks that are likely to be performed mostly in direct day light, or under very bright artificial lighting	Working with no roof overhead during the day
6	Tasks that actually put work in place	Nailing, cutting, digging, and etc.

Situation-aware Interface Design: An Interaction Constraints Model for Finding the Right Interaction for Mobile and Wearable Computer Systems

Christian Bürgy*; James H. Garrett, Jr.*

ABSTRACT

The design of mobile and wearable computing devices involves decisions about how the user interacts with the hardware and software composing the device. Since applications of, and usage environments for, mobile and wearable computers has varied greatly, it has been difficult to build on previously collected design knowledge. The *Interaction Constraints Model* described in this paper offers an application-neutral and domain-independent approach for comparing different applications and usage scenarios. The *Interaction Constraints Model* provides a means to map information about user interface implementations to specific work situations. In this way, a system designer can use a set of generic interaction constraints to identify and retrieve information about user interface components from previous projects. In a proof-of-concept implementation of the *Interaction Constraints Model*, we were able to validate the approach of the model and we demonstrate the usefulness for the design of wearable computer user interactions.

KEYWORDS

Design Constraints; Human-Computer Interaction; Industrial Applications; Mobile Computers; Wearable Computing

1. INTRODUCTION

On construction sites, we see ever changing 'work situations' that differ in their 'work locations' and 'work activities'. Mobile and wearable computer systems can support workers in these *work situations*. But to be useful tools, these systems need to offer specific user interfaces that are appropriate for the *work location* and the *work activity* at hand. The 'Interaction Constraints Model' aids system designers in choosing the right interaction means for specific tasks with respect to the environment in which the task is performed and the kind of mobile or wearable computing system that supports this task.

In this paper, we describe the underlying concept of the *Interaction Constraints Model* and the benefits of using it for interaction design and conclude with showing the result of our proof-of-concept implementation.

2. BACKGROUND

For computer-aided engineering applications, mobile IT support helps to improve construction processes [1] and enables mobile workers to perform their tasks better, faster, and with higher quality, i.e., with higher data consistency (less manual data entry and reentry), shorter data access times (connection to the company's intranet and to online manuals), and better communication means (Internet telephony, short messages, expert forums). However, mobile workers usually perform several different tasks in ever changing environments. This generates different 'constraints' on the system design of the mobile IT support with respect to: the kind of the ***task*** to be performed; the ***application***, for which the task is performed; the influences caused by the ***environment***; the ***device*** chosen as the supporting hardware platform; and the abilities and work patterns of the ***user***.

* Department of Civil and Environmental Engineering, Carnegie Mellon University; Porter Hall 119, 5000 Forbes Ave., Pittsburgh, PA, 15213, USA. Tel.: ++1 (412) 268-5674; e-mail: buergy@cmu.edu; e-mail: garrett@cmu.edu.

Each situation demands that the *user* and the mobile or wearable computer *device* adapt their interaction with respect to the *task*, the *application*, and the *environment*. Many mobile input devices have been developed for use "on the move", such as mobile, body-worn pointing devices and keyboards, scanners, or data gloves. However, these interactions still involve using at least one hand. Some tasks, however, have to be performed using both hands. Thus, it would be helpful to get support in making the decision about which user interface to use for which situation.

3. INTERACTION CONSTRAINTS MODEL

The concept of the *Interaction Constraints Model* is a generic description of different *work situations* based on the *constraints* that impact interaction with mobile and wearable computers in industrial applications. The idea is to compare different *work situations* between different applications of the same or different domains. Thus, we can compare the system design of a mobile IT device and re-use the design decisions. This concept helps to identify similar situations and evaluate how well a specific means of interaction performed in previous applications.

The *Interaction Constraints Model* builds on two definitions: the *constraints* themselves (section 3.1) and the *work situation* that is described by a specific set of *constraints* (section 3.2). A description of the implementation of the model (section 3.3) and a usage example (section 3.4) demonstrate the concept of describing a *work situation* with a set of *constraints*. Section 4 illustrates the contributions and results of the *Interaction Constraints Model*.

3.1 Constraints

Leffingwell and Widrig define constraints as "a restriction on the degree of freedom we have in providing a solution" [2]. Constraints in the *Interaction Constraints Model* restrict the use of a specific interaction modality for a system design for a specific usage situation. Before the introduction of mobile computing, the *constraints* that implied the design of IT systems could be associated with three *constraint categories*: namely "user," "device," and "application." Now there are five categories, since mobile and wearable computers imply changing "tasks" in changing "environments." These *constraint categories* contain *constraints* that influence *constraints* of their own category as well as *constraints* of other categories during operation. For example, a *device constraint*, such as the absence of a display, influences other *device constraints*, such as the need to provide alternative output means; it also influences the *application constraints*, such as no GUI interface being possible. Sections 3.1.1-3.1.5 describe these five *constraints categories* [3].

3.1.1. Task

Tasks are considered to be "states in the working process" as a part of the workflow. *Task constraints* are all those *constraints* that restrict the interaction between the *user* and the *device*, such as a *task* that requires both hands of the *user*.

3.1.2. Environment

Environment constraints are defined as *constraints* of the working / usage *environment* of the *device*, composed of such influences as ambient noise level, lighting, potential hazards (need for gloves, masks, etc.). However, properties of the IT infrastructure are covered by the *device* description.

The considered *environments* are mainly those in which multiple (non-traditional) input modalities are applicable and special demands on the *user* are present (i.e. office environments are covered by existing HCI research and thus not the main target of this research).

3.1.3. Application

Constraints of the *application* influence the user interaction by demanding different navigation

/ operation *tasks* of the software, e.g., a CAD application deals with 2D or even 3D drawing navigation, whereas an inspection application deals more with check lists. There are domain-specific applications, such as construction or manufacturing applications and general applications that for example support the "back office" processes. Furthermore, different application structures or software architectures cause different behaviors of the software. Finally, the *application constraint* category holds the actual interface / interaction layer, i.e. the interface to the *user* of the *device*.

3.1.4. User

User constraints result from different cognitive, logical, and physical abilities of *users*, as well as different expertise and experience. *Users* and their capabilities are also constrained by the working *environment*, such as situations that demand special attention or occupy the *user* in some way.

Working habits of *users* should not primarily go into the *constraints* design, since these habits might change completely with the use of the mobile IT support. However, these habits have to be investigated thoroughly to fully understand the *tasks* that have to be supported by the IT device.

3.1.5. Device

The *device constraints* result from the device itself, as well as from other *IT* components connected to the *device*. These *constraints* are for example the presence of different input / output modalities that are more or less appropriate for a given *task*.

3.2 Work Situations

Work situations are uniquely defined by a combination of a *work location*, the place where the *user* of the mobile or wearable *device* performs a job, and the *work activity*, the actual *task* of the *user*. The following are descriptions of the components that enable the comparison of different *work situations* and thus the re-use of design knowledge.

3.2.1. Work Location

Work locations identify the location, and thus the conditions, in which a *work activity* is performed. The reason to have locations as an identifying factor in the model is the fact that interaction with a device is constrained differently at different locations of one project and at locations of other projects.

Example: "Inspecting a bridge structure" and "assembling tubular steel scaffolding," have many conditions in common; e.g., the sunlight, the height of the workplace, safety concerns, etc. Working on a "tunnel construction project" and in a "pit of an automotive workshop," also have similarities: the artificial light (if any) and the dust / oil of the machines or vehicles.

3.2.2. Work Activity

Work activities represent primary *tasks* of the *user*. Primary *tasks* are the tasks that the envisioned mobile or wearable device will finally support. As mentioned above, *work locations* and *work activities* define unique *work situations*. And the motivation for including the *work activity* as an identifier is similar to that for the *work location*. Here, too, the goal is to find patterns of similar *constraints* that result from different activities and to re-use these patterns for design decisions for new *work situations*.

It may seem hard to compare activities from different domains and to find similar patterns amongst them. But the *work activities* themselves will not be compared, but rather the *constraints* and the *constraints'* influences on the user interaction, which originate in these *work activities*, are compared. Thus, we create *constraints* that are not domain-specific and enable a domain-independent model.

Example: It is obvious that some activities, such as "inspecting bridges" and "inspecting vehicles" have similarities, but even the two activities, "determine the inventory of construction material" and "perform quality assurance at a manufacturing facility," can be mapped to a common constraint pattern.

3.2.3. Work Situation

Work locations and *work activities* define unique *work situations*. The link between the location and the activity is the user who literarily brings the *work activity* to the *work location*. This is a new aspect that is caused by having mobile and wearable computers, which enable IT support away from the desktop or kiosk-like terminals. Thus, we have to identify varying sets of conditions or requirements to which the design has to be adapted, and to which future adaptive devices will adapt automatically.

Each *work situation* is unique in a sense that exactly one *work activity* is performed at one *work location*. However, the conditions at different *work locations* and the demands of different *work activities* can have common patterns, and can thus lead to similar *constraints* on the user interaction with a mobile or wearable device.

Example: We can use the two examples above to show the concept of a work situation: "bridge inspections" and "vehicle inspection" differ in their location; so do "assembling steel scaffolding" and "quality assurance" with respect to the activity. However, "inspecting a bridge's interior structure" and "assuring the product quality in a poorly lit manufacturing plant" have similarities in both respects.

3.2.4 User Interface

Finally, the *Interaction Constraints Model* provides information about user interfaces that were implemented and evaluated in previously conducted projects. The system designer can retrieve this information, which is mapped to specific *work situations*, and use it for designing the user interaction for mobile IT devices for a similar *work situation*.

3.3 Implementation

In order to conduct a proof-of-concept of the *Interaction Constraints Model*, we implemented the model as a database that stores all the necessary information about the *constraints* of *work situations* and the user interfaces that were used in about 15 different previous designs of mobile and wearable computer systems. The implementation allows the user to enter the *constraints* of a new *work situation* and to query the case-base. Figure 1 shows the attributes of the different *constraint categories* that describe each *work situation*.

Each of the attributes of the *constraints* can take several values, e.g. "low," "normal" or "high" ambient lighting. We needed such a simple classification, since the documentation of the investigated projects did not provide more detailed data. However, this classification was sufficient for this proof-of-concept.

As a case-base, we collected project information on 15 system designs of our own research group, from other researchers at Carnegie Mellon University, and from the literature [4]. This case-base was diverse enough to illustrate the usefulness of the application and the domain independence of the model, and showed that we can use *constraints* to retrieve similar situations of previous projects.

3.4 Usage Example

To demonstrate the concept behind the *Interaction Constraints Model*, we want to present a brief example on how the interaction design of a new wearable computer system can be supported by using the model: first, the system designer performs a task analysis and identifies the *work locations* and *the work activities* that occur for the envisioned application. For each relevant combination of *work location* and *work activity* the designer defines a *work situation* and enters estimated or measured *constraints* for each *work situation* in an implementation of the *Interaction Constraints Model*. Depending on the amount of cases entered in the case-base and the query capabilities of the implementation of the model, the designer gets a set of similar *work situations* that occurred in previous projects. Now the designer can retrieve information about the user interfaces used in these *work situations* and evaluate the performance of these user interfaces. Based on that information, the designer can decide which user interface to include in the new system design and which interfaces would not perform well. After collecting user feedback on the design, the designer enters information about the new design and thus adds information to the case-base.

4. RESULTS

Using the proof-of-concept implementation, we performed several different types of tests. First, we took situations that an experienced designer could map without any help to prove that the system is valid; then we used the system to query the database for *work situations* that could not easily be imagined, but served as good design examples.

One example was to compare a progress-monitoring task on a construction site (Progress Monitor by Reinhardt, et al. [5]) with a vehicle inspection performed in the field (VuMan Amphibious Vehicle Inspection System by Smailagic, et al. [6]). The similarity of the *constraint patterns* for the two *work situations* results from the fact that both locations are outside in sunlight, with noisy machinery close by, low cleanliness due to the construction site or vehicle oil, respectively, and rough conditions under which the devices are used for the inspection.

The second example in which the system returned corresponding design examples from a different domain were the transmission of patient data of EMS personnel at a highway accident at night and an inventory maintenance task in a tunnel construction site. The matches derived from the system show that in the different *work situations* of the two applications, the same set of *constraints* restricts the user interaction and thus can be designed in a similar way.

Finally, the system mapped a tourist guide application helping a tourist in a restaurant to find the next attraction in an online multimedia guide and a worker querying a mobile spare part database in a manufacturing application. In this example, the match of the *environment constraints* and the transfer from one domain to the other domain made it unlikely to imagine the match without the help of the *Interaction Constraints Model* implementation. The conditions in an industrial supply room are surely not the same as in a bar or restaurant, but they impact the design of mobile IT system with the same set of *constraints*. Another finding about the tourist guide project is that using the system in a museum restricts the “Linguistic Ability” of the user and the “Audio Input” of the device. Thus, it matches in these categories to many other industrial applications. However, these restrictions do not result from the high ambient noise, which does not occur at a museum, but in the required silence expected from museum visitors, which does not allow for using speech input by the user.

These results showed us that we indeed could compare *work situations* based on the *constraints* that impact the user interaction. However, we have to extend the case-base of mobile and wearable computer design projects to sufficiently support a broader variety of *work situations*.

5. CONCLUSIONS

With the *Interaction Constraints Model*, our approach to determining the best interfaces for a given situation is to map the possible *constraints* for different *work situations* to a set of potential user interfaces. Thus, we map the possible interfaces and their applicability to the *constraints*, independent from the application or domain in which the *constraints* occur. This approach allows a more systematic means by which to search for and apply previous experience gained in mobile computing projects.

6. REFERENCES

[1] Trupp, Torsten. 2001. Mobile Computing in Construction Management. Thesis, Department of Civil Engineering, University of Applied Sciences Giessen-Friedberg, Germany. p. 7

[2] Leffingwell, Dean and Don Widrig. 2000. *Managing Software Requirements – A Unified Approach*. Addison-Wesley. p.44

[3] Bürgy, Christian and James H. Garrett, Jr. 2002. *Wearable Computers: An Interface between Humans and Smart Infrastructure Systems.* Bauen mit Computern 2002, Bonn, Germany. VDI Verlag GmbH, Duesseldorf, Germany.

[4] Bürgy, Christian. 2002. *An Interaction Constraints Model for Mobile and Wearable Computer-Aided Engineering Systems in Industrial Applications*. Thesis, Dept. of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. p. 82

[5] Reinhardt, Jan, James H. Garrett Jr., Raimar J. Scherer. 2000. *The preliminary design of a wearable computer for supporting Construction Progress Monitoring*. International Conference on the Application of Computer Science and Mathematics in Architecture and Civil Engineering (IKM 2000). Weimar, Germany.

[6] Smailagic, Asim. 1998. *An Evaluation of Audio-Centric CMU Wearable Computers.* ACM Journal on Special Topics in Mobile Networking and Applications, Vol. 6, No. 4, 1998, pp. 65-76.

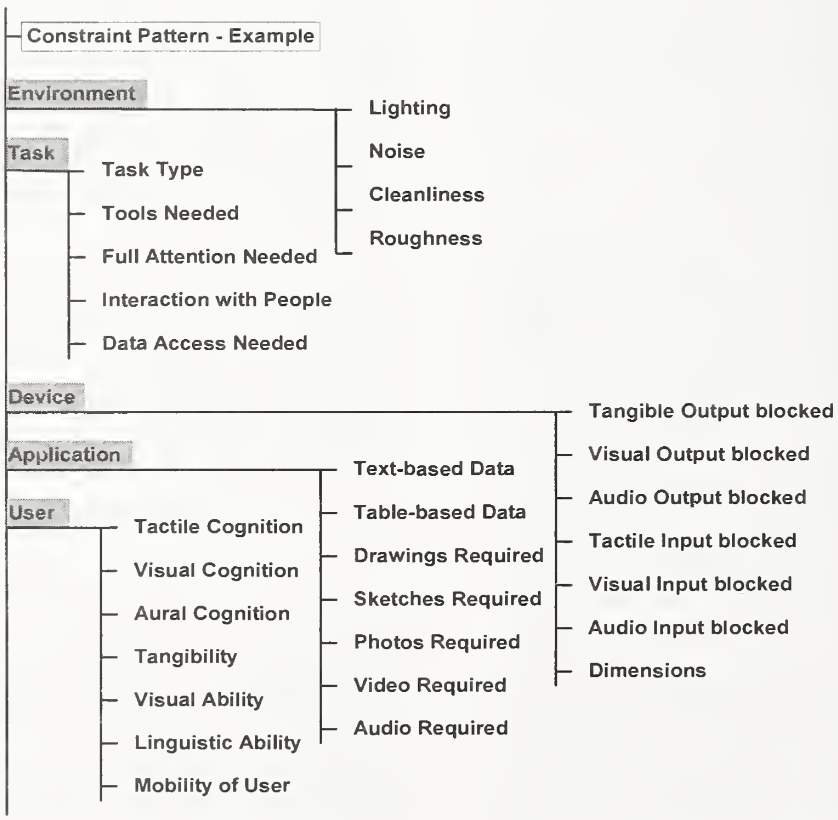


Figure 1: Each *work situation* can be described with a specific set of *constraints* (or *constraint pattern*).

NIST *Technical Publications*

Periodical

Journal of Research of the National Institute of Standards and Technology—Reports NIST research and development in metrology and related fields of physical science, engineering, applied mathematics, statistics, biotechnology, and information technology. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Institute's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Institute's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NIST, NIST annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NIST under the authority of the National Standard Data Act (Public Law 90-396). NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published bimonthly for NIST by the American Institute of Physics (AIP). Subscription orders and renewals are available from AIP, P.O. Box 503284, St. Louis, MO 63150-3284.

Building Science Series—Disseminates technical information developed at the Institute on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NIST under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NIST administers this program in support of the efforts of private-sector standardizing organizations.

Order the **following** *NIST publications—FIPS and NISTIRs—from the National Technical Information Service, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NIST pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NIST Interagency or Internal Reports (NISTIR)—The series includes interim or final reports on work performed by NIST for outside sponsors (both government and nongovernment). In general, initial distribution is handled by the sponsor; public distribution is handled by sales through the National Technical Information Service, Springfield, VA 22161, in hard copy, electronic media, or microfiche form. NISTIR's may also report results of NIST projects of transitory or limited interest, including those that will be published subsequently in more comprehensive form.

U.S. Department of Commerce
National Institute of Standards
and Technology
Gaithersburg, MD 20899-0001